

FIGURE 1A

The Small Island: 10848bp (SEQ ID NO: 1)

TTTTCGAAGGGGCGAGTTGCACAGGCCTAGGAAGGGGTGTGCACCTTTACTTGTCCAGCAGCCGCTCTGGGCGGAACAT
TAGGCCCGCATTGACGTGATGTCAATATGGAATAGCCCGCTATTAGGTGGGCTTTTCTGTTTTTCAGGCCCTGAAACC
ATTCTGCAGGCGTAGCAATGTGCCCTCGACGTAATCAGCGATTACAGAGGGTACTGCGGTTTCGGCCTTGAAATTGCGGG
GCTCTGCCAGCATGCCGCGCTGGAATCCAGGCGCTCTCGGTAATTCAGATTTCCTACAGGCCCTAGCGAACTGCTGG
GAGGCGCAGCTACCTTTCTCGGTAGCCGGCCCCCTCAGTATCGGTATCGTAGGCGAGACTGCTTCGGGCGGTAGATCC
CCATCCTCCAGATAGCGCTTGTGTGCGGTGGAAGCCGCGGAGCCGCTGGCGGCACGGTTCGATAGGTGTGAGCCATCAGAG
TGAACAGCACGCTTTGCTCGAAGGTGTGAGTGGCTGTGGTAGCGCTTCTCCACGCGGTTGCATCTCCCTGAATGCCT
GTCCCTGGTGTGTGCTCCCTTGGTCTGTCAGGTTCTGTCTAGCTCTGTGGCGGCGCTTCACCCATTTCGAGCGAGAGG
TTGCCCGCATGGCTTCGGTTTGGCAGAGGCTTGGCGGGAGCCGCTGTAATCCACCTTCCTGACGTTCTCCAGCGCTCCTT
GGGACCGCGCTTATCGGCTTGTGAGGACGACGCGCATGTGGTCGTAGTTCGGCTCGATGTAGCTCATGGTGGTCTGGATATTC
GAGTGGTTGAGCAGGCACTTCGTGAGGTGATGTTCCGCTCCGCTCCCTTCATCAAGTCGGTGGCCAGGGTGTGCCGGAA
ACGGTGGGAGTTCATCCGACCCCCAACCTTCTCGGTCACTTCGGGTACATGGCTTCGACCTGGTTCGGAGTTCATCACCT
TGCTCTTGTAGTGGGTGAGAACCGGTTGACGTTGAACAACTGGTTCGTATCGGCGAATCCGGCTCTATCGGCCCTCTGTC
AGGAGCCGCGTATAGGTGAGGACAAAGCCCCCTCGGTATTCGGCTACGAACTCTTTATGAGTTTCTCTGTCTCGCGCG
GATGAGGATCAGTTGATTTTCCAGTCGATGTGCGGCTTGGGATGCACAACAGCGCATTCAACCGGATGCGGGTGAAGT
AGAAGACCTCAAACGTCAAAGCCAGAACCAGGCGGGAGTGTATGCGTGCAGATCGCCAGTGCAGCGCTCGCGCCGACC
TGCATGTTGAGCCAATTGCGGCGCGCAGGATGGCTTCGGCTGCGACGGTTTTGCTTGCTCGCCTGGGGGGGATGACGGT
GGTCTTTCTGAACGGGTTGACTTGGGAGTGTGTACCAAGCTCGTGTGATGGCATAGCCCCAGATCGTTCGCGATGAT
TCGAATACGTGTTCAGCTCCGCTTCGACAGGCTTGTTCAGGACCTTGGCGCGCATCCAGCACAGACCTGTGGTCC
ACCTCCTGTACGGTTGCGGTAGGGCGGAAGTGTGAGGAGCGCTTGGTTCGGCGCGCGGTAGATCTTCGCGCTGGCTTC
TCGGAGATCGTGCAGAGATGTAATCTCTCGGTGAGCTGTGCGGCGCTCATTCCACACCTCCTTCGGCATCGGTGATGAC
CCTGAGGCTTGGGTTGTGTCAGAGGCTGCACAGGGAACGCAATTTGGGATCCTGGAGCAGGTAGGCCTTGAGCTCTTCG
TCTTGCAGGACCGAAACCTTGATGTTCCAGATATTAGGTTCTTACTGGTCTTTCGGTGAAGACCTGTTTCTCGAAC
GCGCGCTGCACAGCTTCCAGCCGCGGCTCTCTTGGCTTGGGCCAGCTTTCAACCTCCGGATGCTCCTGGACATAACG
CTTGAAGATTCCTGGCGTGACAGCATGGCGGTCCCGTCGACGGTATGCACCAAGCCTTGGTGTGCTGTATGAACAGGC
GACGGGTCCGATACAGATTTCATCCAGCAATGAATCCCTGTCCAGATCAGTGCTTCTGCTAGGCATATAATCATCT
TCAATTGCTTCAGGAGCGCAATCTGTAGTGTCAACCCATCCCGACAGTGGGTTGAGTTTTTCTTCGGCCACCGCAGGT
GGTAACCCGTCGTTGAATTGATGCGGCTGATCCAGTTGTAGCGGTGCATCAAGTAATGACTGATGTCCCGTTGGGCGCTC
CTGCGCGCTAGGTAAACCGTTGACGGGACCACTCCGACTTCAGACTGCGGAACAGGCGCTCCATCGCGAGTTATCC
AGCAATTCGCCGACGGCTCATGCTCTGCTGCTACCGCATAGCGCCAGAGCCGTTGCCGAACAGGCGGCTGGCGTACTGG
CTGCCCTGGTCTGAATGGAACAGCAGCTGTGTGGCCTGCCGCGCTGTCTGAGGCCATGTCCAGGCGCTTGTATCACCAG
TTCGGCATCCGGCTTGGCCGAGAACGCCAGCCGATCAGCAGATCCAGCAGCGCGGCGAGGTAGTCCAAACGCGCTTGGC
CCAGAGCTAGGTGATGTGCGCACACCACTGTATGGGATGCTCGGTGCGGAATTCGCGGTTACGCCGATTTCGGGATA
TCCGGCGCTCAACCGTGGCTGTTTGTAGGCGTGCAGGCCCGGTGCTTGTGCTGACAGGCCAGCTCACGCATCAACCG
ACGCATCCGGAACGCGCGATGGTCAAGCCCTCTTCGGCGAGCATGCCAGAAATGTGCGGCTGCCGCGGAGCCCCGAC
TCTGGCTGAACAACTGGTTGACTTGGCTGCGTAGCGCCAGCGGCGAGCATCGACAGCGCGCGCTCGAAGACGGTGGACG
TAGTAGCAAGACCGCGCCCATCGAAAGCTGAGCAGACCACTTCACCGCACTCTGCTCACTCAACTGGTCTATCAGCGC
GTACGATCGAGTTCGTCGACATCAAGAGAGCGGTAGCCTTTTAAATATCGCTTTCTCCGCTCCAACTCGGTTGATCCG
GGCTTCCAGCTCCTGGATCTTTGTGCTCAGGCGTCAAGCCTTGTCTTTCGGGGTCAACCCCTGGCGCTCCGCTCGA
GCTGCTTACCCCAACGCGCGCAAGGCCGAATCCACACCCCCAGCGAACGGCAGGCGTCGATATGGCTGTAGCTTGGTCC
AACACAGGCGCGCGGCTCTCGTTTGAATCGCGCGGAACGTAACGTCGTTGCTTGTCTCATCAGACACTCTTTCACGG
CGAGGATTCTCGCTAAATCGGTGTCCGGGATCAGTAGACCACTACACTGCCGAAACAGTTATTTTGTGCTAAGAGGC
AGGAAGGCTGCGCTGAGATTCTACTAGTGAATTAGATAGTTGTTGTCTTCTGAAACAGAACTGAAGCGAAATTGGGGGT
AGGGTTTTCTAGGTGAAGGTAACCTTTGAGAATTACAAAGGTGTGACGCAATGTTGTATTTTCTTGCAGTATGAAG
TGAAGTGTGAGGTTGGGTGAGATATAGTACTTCTCTCTTAAATTTGCTCTCATCTATGGGTGTGTGCGGTGAGG
TGGATCGGATGAGATTGGGCAGCACTGCTTGAAGAGAGCAAAAGCTTTCGGAGTTAATGATAATGAAGAGGGGAGTG
TGAGGTTGAATCGGCTGAATGCGATCCAATTGAAGTTCGTGTTCTGAATCAGAGAAGCTGATAAGAAAGCCGCCAAT
GAGCTGGGTATTCACTGAGTTAATAATTTTAGCCTGGAGAGGTCTTTTGAATAAGAAAAATCACTTGTATGTGCGG
TGTGCTTTTGGTGGCTAGCAATTTGCGGTGTGCTGATGAGGCTCAATGATGGAAGTGAATATGTGCGGCGCAGGGTG
GAGTTGAATAACAAGTCTGGGGGAAGTCTCAAGGGGTGTGGATGTTGAAGATGTTGTAGTTTGTTCGATTCTTCAAGT
AATATGAAGTCGAGTCAAAGAGCGCTACACTCCCTCTCTGCAAAGGATGATCAATTTCGGCAATGCCTTCACAGGAAC
GGTCACTGTTTCTGCCAGCGGAGATAGGAAATTTACAACATCTTGCAGGCAATCTTTATGCTCCAGCTTATGCCAAT
TCTATCCAGACGGTGTAGCAGGGGAACATCAGATCTACGATGTGTTGGTTACAATACACCGGGAATTCATCTCAAGG
TGTAATGTGTCATGGGACGCGCCGACCGACATTCAATTGGGTGTGAGCCATATGGCGGATCTGTTGTTGTAACCTACAG
TTGCACTGCATTCAAACACAGATTCCAGTGATAATGAGCTACAGTTATCGTGATGGCGGGCAGTGATGCGCAGGTCC
AGAATGTGTGAGGAATAATAATGTGGTTTGAACAAAGGATGCGCCTTAGTAGGATTTTATTTCAAGGTGTTCAAGCTTG
TGAATGTGTCAGGCACTCTATAGAGTGCCTGCTGGAAGTTCGTTGGCTGAGGGCTAGTCTTTCGAGTTG
ATATGGCACGTTGATGAAGGCTGGTAAAGTCTGTCTGTGTTGGGTGGCAGGCGAGTGGCGGCTTACAGCGCTAATGTG
GAGTTCAAGGTAGGTTCCGACGCGGAGGATCTCTTTGGGTGGAGTCAATTGCTGAGCCATGCAATTATAGACAGGCT
ACGTGCTCCATTGACTTGGACATCGATACAGGAAGATTCAAGTTGTTCCGAGTGTGACTCTGCGGGCGGGGAGGAGCAGCAG
CCTGCTGTGCGCAACCACTATAGGCAACAAGAAATATAATCGAAGAAATTTAATAAGCAATCCAATCTCCTTAACCTG
AGGGTTGGATGAGACTCAATAGTAACCTCAAGTTATGAGGAAAGATGAACCGCTATCGTTTTGTAGTCTTGGTTAGGTC
AGGCACAGCTAGTGAAGTACTAAGAGTAAGCCGATAGCCAGGGGCGTGGTGGCCATTGGGGGAAGTATGAAGAG
GTGCGGCTTGTACCGGTTGTCAGGAAGGAATAGACAGCAGACAGGATCGCTTGGATGGGGCAGCAGGAACCATCAGCC
CGTCCGATCCAGCACTGCTGCTGTCAGCGTTAGGGGTGACTCCCTCATGCCCTAGAGGGCATAATCCCTGTATGAGAT
GGCTAATGATATTAAGGAGTGGTCAAGTTCATGGAACGCTTGTCTGAGAGCAATTTACATCAATGCCCGCGCGCGATGG
AGTTGAGGCTTAGCTCACCAGCTCCGCGCGCAAGAGAATGGTAAAGATTGTGGATGGGGAGGAGTTCGAGGTTCTGCCA

FIGURE 1B

GTGTTGAAGTGCAGGGCATCTCTGGAGGCCAAAAGAGGGATGTTTGAATCCTCGCCGACTTCTTAGCCAAAGAGTCTCTGGTC
GCGACGCTAGAGATCTTGGCTAGCCAGTCTATCGCCGACAGCAACGTGGTCTCGTCGGCCATGCTTCAAGCTTACAGCG
TTCGTCCTGCTCTGGTTCCCGGATGACGAGGGTCCGGGTCGCTGTAAATGCCCTTCGACTGTGCGGTGCTGATCTCGTA
AGCGTCCGGGAAGGCACATTGCGATGCGCTAAGCTGTAATCCTTAGAGTAGAGGTTCTGACATGGAATGCCAGCTGCG
TCCCGCCACGAGCAGAGATGACGACGCGATAAGCTGCGTAGTTATAGCCGCCCTGCGTGAGTCAAATTCACAGGACTATC
CGCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTTCTCCTGAAGCCATACACACACAGCTTACGAAAGCGTAGGGTCTTC
GTAGCTTATTTGGGCGAAAACATTATGGCATGCGCGGTCTGACGGCTGACGTCGTGAGAGTGTCTTTCGTTGACCCAGC
TACACAGAAAGCGGTATCGGGCGCATTTGATGGATGTCATTACATACAACCTGCGCAGCGGGGATGGAGCTGTAC
GTGTGCCATCGTCGATTACAGTGAAGGTTTTATACCGCATTTGGGTTATCAGAAAATCCGCGACGAGTTTCATGGGGCG
GAGCGACCACTCGTTATGGAAGAAGCGGCTGTAGGATTAATCCAGACTTAATACCCACTTATGCGGCAACAAATCAGTGAT
TCTACTGGCCCGCTGCGTGGGCGACCGCTGACACATCTTTAGATAGGCATACGAGTACGCTTATCATCGCCGCGC
ACTAGATCAGGCTCATGATTGACGCGCCGCTTTACCTACGCGACGACCGGCAACAGCCAGTTGCAAGCCGCGGAGC
GCCATGCGCGTATTGGTTCTCGACCTGATTGTTATCGATGGGCACAGCCCCATCGTCCAGGTAGCGCGTCAGCGCTA
CCAGCGTTTCAGGCTGTAAATCGAGGGCTTTGGCCGTGGCTGATCCATTGGGCACACAGCTCTCGCTGAGCCAACTCCAGT
CATGCACTTTTGGAGTACGGCACCAGCAATCTTGTGCTATTGCCAGCGCTCTTCTACTCATGTCTCCGCGCGCTGA
CGTTGCAGCTCGTACAAGCCGCTGATCGAGTGCAGGCTGTTTACGCGAGCTGACTTTTGTTCGCCAGCTGCAAAATCAA
GAACTTGCGCGGGGCTGGGCCATGACGCGGATTTAGTGATGCTTGTTCGAAACCGGCTTTGTAGCCAGCGAAGTCGT
CGCAGTACGCTTCCGCTCCAGTACCCAGGAAGTTGCGCGCATGTTGCCAGCAGCGGCTGGGCTGAAGTCGTAACACC
ACCGTTTGTAGCCCTGAAACCGCCGCTGCTGTACGCCAGACATAGGCCGCGTGGTTTCTTCTCGCTGGCGCAAG
CATTGACCGGCTGTTTCATCAGCGTGGATACGCGCTGGTTACAGCGGCTTACGAGTGCATGCACGATGGCTGAA
GCCGACGCGCGTTTGTTCGACCACTGCGCCAGGGTCGACGAGCAATTGCCAGCCGCGCGCGGCAAGATTTTCTCC
TGCCGGTACAGCGGCAAGTATCGGCAACTTGGCCACCATCAGCTGGGCGCAACAAACCTGCGGTGGGATACCTTTGTC
GATTGACCTGCGCCCGATTTAGTACCACTGGAATCTAGTAGAATCGTTTTCAAGCAGAGACGGCAGTGCAGAG
CGTTCTTCTCGACTTCTCAAGCGGGTCGAGGCTAGTGTACCGACAGAGCTGCACGCCCACTGGTGGAACTGGCCACAG
AGTATCGACGCTGTAGCTAACGCCGCTGCACATCTGCTGTGGCGTGCAGGTCAACCAAGCGGATTTACCGGCTGTA
CCGACGCCGCGGCTGATGTTGAGGCGCGGCGGGTCCGGTGGCGCGCAATGCGCTGAGCTGTGACGAGCCACCGAAC
AGGCTTGTGCTAGGATTTGCTCTTCGACGCGTACAGCTGGCGAGCGATCAATGCTGACGTTGGTGCATGACTTCT
ACCAAGGTGTGCGTCGACATCTTGGTGGAGTACGTTATCAGCGGTTTTCTGTGTACGCGGGCGCTGGACGAGATGGCGG
GTTTCGTGGCTACCCGACGGCGATCCGACCGACAGGGCGCCGAGTTACCGGCAAGGCGCTTGATCAGTGGGCGTGTCT
AGCGTGACATCAAGTTGAAGCTGATTACGCTGGCCAGCCACGAGAGCGCTTCTCGAGTCAATCAACGGCAAGTTCT
CGGGCGAATGCTCAATGAGCATGCTGCTGCTGCGAAGCCAGAATCCGTTACGCGCTGGCGGATTAACACGAGCA
CCGACCAACACAGCGCCATTGGCAATCTCTCCCGCGAGAGCTTGTGCGAAGTGGCGAACAACACGACGAGCTGAAGC
GGGAAAAAGTTGATATCAACCCCTAGCTTACTAATAGCAGCGGCTACTAAACTGGGGCGAGGCCAGTCCGAATGATA
AAAAACGTCGCAAACTGGCAAAACCTTATGGCCGTTTTCTAATATTGCTCCGAACCTCGGTATTTCAAGGAGCGA
ACATGCATATCAATCGTTGGGGGCTAGCTGCTCTCGCTGAATCAGAGCCGCTGCGAAACCCGCTCGAGGACGCGCAT
AAGTCCGCCAGCTTGGCTCAGGAACCTTCAAGGCAAGGTCTCGGGGTTGCCCTAAGAGCACGCGCGGAATACTTTCCGG
GAAGTTCGCGGAAGCGTTAGCGAGCTCGGTTTACAGAGTCCCAAGGGCAAGGGAGTCCCGTACTCTGACTGACTCGG
CAGGCGCGCGGACATCACTCTGCGCCAGTTTGAGAACGGAGTACCGAGCTACGCTCAGCTCGGCCACCATTGACAGT
CTGGTCTTAAGCGGCGGTGGTGCCAAAGGTGCGGCATACCCGGGAGCAATGCTGGCGCTAGAGAAGAAAGGATCTCGA
TGGCATCCGCAGCATGTCGCGTTCTGTCGCTGGCGGCATACCGCGCGCCTTTTGGCTCAGGTATGAGCCCGCGCGCT
TCAAGACCTTTTCCGACAAGATGGATCTTATTTGCTGCTCGACGCTCGCAACAAGAGCTGAAGCTGTTTCCAACACTT
AGACGCGAGATCGCGCATCTGCTGAAAAGGGCTTGGGCAACAAGATCGCGCGCTTCTGAGTTGTGCTCAATGTACT
CCCACGCATAGATTGCGGGGCTGAGCCCTAGAACGCTTATTGCGCGACGAGACACGCAAGGCCGCTGCTCGGACAGATCG
CTACGCATCCAGAGGTTGACGCGCAGCGACCGTGTGCGCCATCGCCAGCAGATGTCAGTCCGGCTCCGAGTACACTT
GGCAGCTTAGATCGGTTGAGTGCTTACATTTCCCAAGATTAAGCAGTGAACATACAGGATACGGCATGCTTCGAGGGCG
TCCGCAATTAGTGTGTTTCAATGCCAGGCACACACCGGATCTGGAGGTCGCCAGGCGGCACATATCTCGGTTCTTCC
CAGGAGTGTTCAGAAGGTGAGTGTGATCAGCCGTACCAGGCCGCGGTAGAGTGGACAGAATTCAGGATGGCGGG
GTGATGATTAACTGCGCGTCCCTGAGATGATGCACAAGAATTTGACAGCGGGCCATCGCGCGCAACGACAACTGAT
CCTTGAGTTCGAGGGCGAAGCTGGGAGGTAGCGCCGACGAGGTACTAGGGGCGCGCGCTCAAGGGCTGGGTCTGCG
GGTGCTGCTGCGCGCGCGGAAATGCTGCAGCTCAGGGCGCTGAGGAATTCGCGAGCAACAGGCTGTGGTGCGG
TTGAAGAGCGAGCGGCTGATTTCAGTGGCATGCTCGGTGGCACCTTGAACCTTCCATGCGCGACGAGATCAAGGCGCA
TCTTCAGGAGCGCTTCAGGAGCGAGTCGGTGAACATCTGGAGAAACGCTTTCAGGCTTCAGAGCGTACATCTTCGTT
TCTCGACGAGGCGCTGCTGGCATCTGATGACGATGCTCACCAGTGTGCTCAACAGAACCCGAGATACAGACGGG
GCGGTGGCTTTTTCGCGAAGGCGCGGATGCGTTACCGAGCTGACTGTGCTGATCTGTTAGCGCAATGGCTTGGCGG
TAGGCTCAAGTTGACGAGGCTATGCGTCCGCTCTTACGCGACTCGATGCGCTGGCAGATACTCCGAAACGCTAGCAT
GGTTGCGAGCTGAGTTGAACATGCTGATAACGTTGATCATCAGAGTTACTCGATGCCATCGCGGGCAGACGGTGCAG
TCCGCGGTGCTCGCGCTCGGTTAGCAGAGCGAGATCGCCGCAAGTGGCGTTATTGCCGGAACATTGTAAGGAAGT
TATCTTCCCTCTCTGTATCGCCCTGGCCAGCGGATTTCAACGTAGCTGTGTTACGTCGGGCGAGGACGAGCTACCG
ATGCCACAGTCCGGCGGAAATCAATCAAGCGTGAACGATATCGTCGACAATACTCGGCACGAGGCTTCTGCGTTTC
GGCAAAACGTTGAGTTGCTACTACCGTTGAGATGGTATAGGCTTGGCGGAATAAGGAGTTACAGATGATTGATATGGCTG
GCACAGTGGGCTGTAGACTTCTTCGAGCAACGATGCACGTTGCGCTGCAACCGGACAGGGACCGGAACCTGTTAT
GGAGCGCTCTGAGGGCGGTTGGCTTTTCGTCGTCAGTGTGGGACTTGTGCTTACGGTTACCGGTGCGGTGCTGATG
AATTGTTACAAGTGAACCTCCATCTCATCTTGGCACCGGTGAAACTTGCAGCGGACGATGCGGTTAGACTGTGCTC
TGGGCTGAGGACGCTGATGGCGTTGACGATGTGATGTCATGAACCGCTTGCACGATAGGCTGCGGGAAGGACATTCAG
ATTAGTGCATTGTAGAGCCACGGGTGAGTTGGTTCCAGTTCAGATCAAAACAGCGGCTAGTGTTCGTTTGAACGC
AAGGGGAAGGAGCGATATGCCAATAAAATCCGCAAGTTGGTAATTTCTCTAATAGATGGGCCACCGAAGGTGGCCTATC
TGTTATGCTGCAATGAGTTGTGCTGCCAACACCATTTCTACCGAGCCCCATCTGGTTCAACGAGCCGAGGCCCTCGAG
GTGGGCTTTTCTGTTCTGGAGTCCGATATCTGATCCACAGCGCGGCAAGACGAGCGGCTGGTTTCGTGCTTTCCGAC

FIGURE 2A

The Big Island: 84830bp (SEQ ID NO: 2)

CACCGGCGATGCATGGGATGCCGTAGGGGTGCTGATCCAGCCGAGGCCGCTGGATGGTTTCTGGTCCGCGCTGGCG
GTCTGGTTGTTGTTGCTCTGGTTGTCATTACCTCCTCGGCCGGCGCCGGGAAGGTGCGCACGGTCCCGTCGTTGAACAC
GAACGTGAGGCTGCGGATCTGCCACGCACGAGGAGGTGTCAGTCCGCCGAGGCGGTCCCGTGGCTACCGCGCCGG
CGACGTCCGGCAGCTCGATGCCGTGCGGTGAGGTTGTCCGGGCGATGAGGATTTTGAACGGGTAAAGGATCATTGACC
GTCCCGTCCGACCGGCACAGACCGATCAGCGCAAACATGGCCACCAGGCCCATGAGCGTCGAGTTCTGCGGCGAGCGTGTA
GGTCTTACGGACGAGCTTGGGTTTCCAGGTGAGATCGCTGTTGGCCGACGGGTGCAGCCCGTCATCGATCCGCTCCA
GCGGTTCTGTCCGCGATCGACCGCATTGCCGAAGGAGGTGCGGAAGCTGAATCCGCTCGGTTGGGTGGTGGAGCCGCG
GCCAGCGGTGGCCATTGGCATCAACCGCGCGGGCGTCTCGGGGCTCGATCCAGACGATATCCGATGAAGACGAGCCCGC
TCCCTGAAAGTGCTGGCCATCCTTGGTCCGACGCCGAATCTACGGGCGAGTCCGTTGTCACCGGCTCCGGGAATGTGGG
TAGGTTCTCGAGCCGTTTCTGTACCTGGTCCAGGATCGTCTGGCTACGGTTTCTCTGCTGTGTGCTGAGTTCTGGGCG
GTGTTGTTGAGCTTCTGCTCGATGTTCTGATCGATGTTGCGCAGGCGGCCCTGCAGGTTCTCATTGGCGGCTTTGAGCGA
GTCATTCTCCTGGATCACCTTGTGATCTGGTCTTGGAGTGCAGGCTTTCGCCACGATGGTGGTAGTGTGTGCGCGG
GCGTGTGCGCGTTCGATGCCAGGGTTGCCGCTTCTTCCGACGTTACCGTCCGGGTTGCGGCGCCTTGTGACTGCGCACTT
TCTTTCTTCCGACGATCGCTCAGCGGATCAGGATGGCGCCGATCAGCAGGGGACGACCGCAATTTAGGAGGGGATT
ACCGGTGAGCGCCACGACCGGCTCCTTGGGATCGATCTGGCTGATGGAGGAGGGCAGGAGCGCTCGGCAAGGCCCGG
CCGCGCGTCACAGATACAGGTTAGTGGTGTGCGAGGCGTCCGCCCGGGCCCCAAGTACGGGTGCTGGAAGGTGCGCGC
GACGAAATGCCCCATCAGTCCCTGGGATCCAGGGCCAGGTGCTGGGCGCTGCGGTTCTGACGCTTACCGCGCTGATGT
AGTAGTCTCGACCGCCAGGCGCCAGGCGGTAGGCGGATGAGGTAGGCTGGGAGCAGGGTGGTCAAGTCCAGCTGT
CGCTTGACGCGCACCTGACCGACGCCATCCACCGGTTCCACCGTGCAGCGGGGATAGAGCATCTGCGCCGATAGCG
CGTCAGAACACGGGGACGGGCGTTTCCGCGCGCACGGCCTTCGGTGTCTTCTGCGTGTCTGGTCTGTTTCTGCTGCTGCCG
ATGGCTGGGCTTCCCGGACTGGCCATAATCGGGATCCACTGGCTCGCCGGCGACGATCCTGACCGGCTCGCGCGGCTGT
TGGTCGGCGGTTGCTTCCGTGGCGGCGATATCGATGAGCATCTGCTCGCATTTGGTCCGCTCTGTAGGCGCAGGCGCGC
TGGAGGAATCGGCTCGTTGGCGAGCAGGTAGAGTGCAGCGCGGCTACTCTGACGCGCAGCTTGGCTGACAGATCCCGAG
GAACCCAACTCGCAGCTTCTGTGACGAAACAAATGCGTTCCTGGCCGACCGTCAATGGAATGGCCAACGGAATGCGC
TCCAGCGCAGAATCTCCACCGCTGGGCGAGTGTGGTAGGGCAAGCATTAGCAAGAGCGAGCCTGTGACTTCCGGAT
CATAGACCTCCCTCCCGCTTGGTGGGCGCGCGGCTCCTCAAGCTCGATACGTTGAGGCGTGTGCGAGTAGCAGTCCCA
CTGACGGCCGAAGGGATTGTTTCCGGGTCGACGTGCGCGCGGATAACGTGCAACGGGTAGCGGGCCAGCGCCGCTTGA
TCTTCTCGCGCGGTAATACTCCTGCTGTCTCATGTCCAAGTTGACGGTCCAGTCATTGATCGAGTGTGATCACGCGG
CCATTGCTCTCGCAATGCTCGACCGGGGATTCCGAGGTGGTGGCTCGCGACCCCTGAGCTCGCGCGGCTTACGACG
AAACTCAAAGTCTTTCTCCAGGAAGACTTTGACAGGAGGAGTGAAGTAGGCGAGCGTAGCGGAACAGGTTTCCCTTGTAGT
CCACTCGCGCTCCTTGGGCCAACGCTGCACCTGCTGGAAGATGTAGAGGCCGAACGCGATAGACGCTCTCTGGCGGAATG
TCCACCACAACCGGGTCTTCTGAGCGCAAGTCGGGCGGAATGTGGATCCAGAGGTCCTTACGGGCTTCCAGACGCA
GTAGGCCAGCACAAAGATGACCATGCACGGAAGCCGCTGATGAACCGGAAGGTGTTGATGTGTGCTGTGTTGCGCAG
TGTGTTTTCTGAAACTCATGGGACTCCGTCGCTCGATCGCTCAGGCGCGGAGCGCTAACAGGCGCAGGCTTCCCGA
AGACGGGAAAGCGCAGCGAGAGGGCCATCTCCACTGGCGGTAGAACAGGTATCGGGCCGCCCCGCTTCATCCGCGC
AGGACGCGGCTGCCGATACCTAGGCGGAGGGCGCGACAGCAGCGCGCCAGTGAATGCAGGCGAGCTTACCTGCGAC
CAAGGCGCGCGGATGCAAGAACGAAACCGGCTGCTCCGCTGGTGAAGACCGTGTATCCACATTTCTGCTGCGGTGAGG
CGCGATGACTACCGGTTGCCGTTCAAACGGGTGCGCGGAGGAGCTGAGGGTTCCATCCTGAAACAGATGTTCTTCCGGG
ATGGGAAGGCCCTTACATGATGGCGGTGGCTTTGGTGACGAGATAAATGATCAAGATCAGCAGGCCGACACCTACGGCTA
CGCCCGCTCCGAGATCCGACCACTTCTTCTCCGCTCATGGATGGCGTGTAGGTGCGGTAGGTATGCCAAGCGACCCCC
AGAAAGACAGCGCGCAGATGAGCAGCGGAGGAGCATCGCTCATAGCCGAAGTTCTGGATGGTTTGATGATGCTG
CGATCCCTCCCCACGGCTAGGTGCTCGGTTTGGGAGTGTCAGCCAAAGGCGATACGGGAGCGCCAAAGTGTGCGG
CCAGGCTCTGGCAGAGGGCGGACAGTTTCTGGAGGTGAGTTTACGATGAGTGTCTCCGTATGGGTGAGTGGAGGGA
AGAAAGAGGTGATGCCGAGGAGGACCAAGATCCGGGATCGCGGAAGCGCCACCGGTGGCGCTGTGCAAGGTTGTTGGTGG
CCAGCGCGCGCAGGTGCTGATCATCGCCAGGCGGACCAAGCAGAGAAGGAGGTGATCGCTGCTCAATGAACAGTCCC
TCGCCGCGGATGGGGGAAAGCCAGCGCGGCGCTGGAACGCTGATGTCTGGGCTCCGCTCATGCTCATGGCATCGGCCCT
CCGCGCAGGGTGAATTCCTGATAGTTCGAAGGGTCAAGCGGTTGGGCGCGGCTGGGCGTCATGATCTGCGAGGCC
TTGGCGGATGCGCTGACGGTCAGCGGCCAGGCGCGGTTAGTTCGAAGTGAAGCGCTGTCCGGGCTCATCGCGCCCTGGG
CACTGCGCGGGCGGTTCTCGAGGGCGTTGAGTGCAGGATCATCACTCCAGGTTCCGCTGCTCGGAGGCGCTCGCT
GCGAAAGCTCCTGTGCGACAGCGCACTGCAGGCCAAGACGGCCAGGCGAGCTTCGGCGAAGGGATGGATGTGGAGTGT
CAGGGCGAATCTCGTCGAGATCAGGTCGATCAGCTTGCCGATCGCCCTGATTCCATTCCGCAAGAAATACTCGGCGCG
GGTCCGGGTATTTTTATGAGGTTGCGCGCTTGAAGAACAGCAGGTAGCTGACGGGCGAGCGAGCCCAAGCCAG
CCTTGGGCGTTTTGATTACCGATCTCGGTTACGCCCGCATGAGTGGCTGAACGCTACCAGCTCCAGGCTTCTGGGCT
CAGACGCACGACCTCGGCCAGCGCAGAGGGTTGAGCAAGGCGTGGTGTGCGGTTCCAAGGTGTCTTCTTGTGTTGA
ACATCGCGACGCCCCTTGGTGAGGGTTTGTAGGGAATGACGATGTTCTGACATGGTATTGCAAGGGCTGTTTTATGCTC
ATGGTGTGCTCCTGGATCGGTTGAGTAGAACCGCACGGTTTCAAGAGCAGGAGGAGGGCTCAACCGGAAATACTAGGC
CTAGCCAGTGAATTTCTGGTGGTGGCGGTTTGAAGACAGTATGGAGCTCTGGGTGATATTTTATATATATTTCT
AAGGGAGACGGCTGATGCTTAGAAACATCTCTATTGGAGTTTGTAGCCATGGCTGTATGTTGGGAGTGTGGGCTG
GCTGCCCTACATTACGATCGGGTCCGCAATTGTTAGTGAGGCGACTTGATTGATGATGTGCTTAGAAAGTGCGGCAA
CCCTGATAGCCGTAATAATGAAGGGCCCGCAGTGGATGGTGGTATAGTGCAGGGGGCTGCTACTGTGCAAACT
GGGTATATGGACCAAGGAATGGATGGTACCAGAAGCTTAGGTTGTGCTAGGAAGACTAGTTGAGATAAAGGCGAGTATG
GACTAGGATAGCCGTGGATGGTGTGTTTTATCCACGGCTATAAGTCTCATCCGCGAGATGATATAAGGGTAAGGATA
TTTGCGATTGTAGGCTTGTGCGTCGGAAATAAACAGGTTGACTGGCGCTGCACGGGAAACTATCTGTTGTAGGT
TGTTCCGATAGACATGCCACGTTGTAATTGGCTTGAATGCTTGTGGAACCTGCTATGTCTAGAAGTTCAAGTAGA
GTGCTTTGGTTGGCGGATGTCGGAGTTGATGCTGATGTTGATGCTGCTCGAGCGGCTGCTGGATGTT
GTTTGAAGTGAGTTGGGCACCATTAAGTTGGCGCTGAATTTTTTAGTGTGCCCGCATGTTGATGGCCTGTAAGGCG

Figure 2

FIGURE 2B

TGGATATATTGCGCCGCTGAGCGGTGTGACGCTGATATTGGTACCCGAGGCCGAGCGCTGATCGCTTGTGAAGGTTG
TCTGCTGTAGTTGTGACCACTTGAAGTCGGCGCTAACTTTTTCGTGCTTTTGGCAGAGTTGATGAGGCTAGTAGGGT
GCTTATATTAAACCGCTGTGCGGTATTGACGCTAATGGATGTTGCGGTTCTGCGCGCTCTAAAGCTTGTGAATGTTGT
TTGAAGTAAGTTGAGCGCCATTGAACCTCTGCGCTGAATGTTTGTGTACCTGCTGCATGAATAGTTTGAAGAAGGGCG
GTTATATTGCGCGTTGTGCGGTATTGACGCTTATGCTGGTGTGTTGTTCTGCGCGATTACTGCTCTAAGTAGGTTGCT
GCTGCTGAGTTGGGCTCCATTGAATCTGCGCTAAATCTTTGAGTTGCGCTGCTGTAGATATTGCGGAAAGAAGTGTGT
TTATATTAAATGGCTGTGCGGTATTGTTGATACTCTGTTAGTGATCTGCTGAGTTTATGACTTGCTGAAGGTTGTTA
GCAGTCAGCTCTGCGCGTTGAAAGTTGTTGAGAGGTTTGTGTTTTCTGCACTGCTGTTGCTTGCCTTGAGAAGAGTTTC
TATATTTAATGCTTGCGCGTTGCTTATGCTGATGCTTGTATTTTTTCTGCTTTTGTAGAGGCTCTAGCATGTTGCTTG
GATTTAGAAGGCTCGCTGCAATCAAGGATAAGTTTCACTCTCAATGCTCCGCTGTGACCTAACGTTAAAGTCTTGGTGT
GAGTACTTCTAGGCTGAATGAGTTGGCTACAAATCTCTTTCGCTTATGTCAATATCAGGCATTGCTGTGACCAATATC
ATTTTTGAGTCATTTTTTACAGCTAGCTATCTGCTCCTTCAGTTTCTGCTGCTCCAGTACCGCTGCATGATCGACAGC
ACCAAGGCTGTGATGGAAGGATCTTCACTGCGCGGGTCACTGCGACGTAGATCAGGTTCAACTCATCGTCGCGCTTGC
TGGGTGCGGTGTGCGGGGCCAGCGGGTCCGCGTTGAAGTCTGCTGACGCAACGAAATCCCATCCAGCCCTTGGCT
TGTGTCGCGTGGTCAAGGTTGTTGCGTCCAGCTCATCGTCAAGGTCAGTGAGCGAAGCTCAAGGATCCGCGCAGGC
AGATCAGGGTAGGTCGATATGATCTTGTATCGAGCGAAGCATCTCACCCTCTGGCTGATCTCGCGATCTCCACGTACTG
GGTGTAGTCGCGGTAGTCACGAGCAGTTTCTGTTCTGGACGTTTTGGCGCAGGCTCGGCTGAATGCGTACAGATCTCT
CCAGGTCGCGCAGCGAGTAATGTCGATACCGCTACCCAGTGGAATTTGGGCTCCGATGATTGCGGACAGCTGCAGG
GCATTTCTGATGACGCGGATAACGGTGGGTGAATGAAGTGGGTGAGGAAGGTCGCGCGGAGGACTTTTTTACCAG
CGTCTGCGGACCCAGTCTTGAAGTTTCCGTGTTTCCGCTTGTAGGAGAGGATGATGTTGGCCACGTGTGCGATCGCGG
GGCGAATCGCCAGCTCTGGGTCAAGTAGTGTCTCTGCGCGCGCCATCCAGTCGCTGTTCAAGGCTATCTTCTGCGCCC
CTGAACCGGTAGAGCTGCTGATGGGATCGCCGACGATAGCCATTCTGATGCGCTGCCAATGGGCAATGTCGCGATCAC
TGGGTGATGCTCTGCCCCCTGCTCAGGAGCATGACGTGCAAGCGCTGGCTCAAATCGGCTTGTCTGATGATACAGCT
TCAGGTAGCCGTCAAGGGGATCAGCATGCGGTGCTGAGATCGACCATGCGCTCCAGCTACTCGCGCCATGTCC
AGGCCCTGCTTGAGGAAGCGTTCTGAGCACTGGTGAGGAACGCTTGTGCGGAAGCGCGGGAAGTGGGTGCGCCGAG
TTCCGCTGCGCGCTGGCCATGTAGTTGTTTCAAGGTGGCCAGCACGTACGTACCAACTCCAGTCTTGGGTATCGAGTC
CGCGGGGATATCGGTGAGTCGAGGTTCTTCTGTTCTTGTGGGCTACTGAATGCCGTACACCGCATGAGCCAGACTG
TGGGCGGTCTTGACACTACGTTGCGGGGAACTTGCCTTCCGCGCTTCTCCACGAGCTGTTGTAGCAGAGATAGAG
GATTTCTCAGGTTAGGTTGACCTGGCAAGCCACCAGGTTGGTATTTTGCAGTGCCTGCGAAGGCTCGCACAGGA
TCTTCCGTGCTTCCAGTGGATGATCGGTGACTGTTCTGAGTCCACTGACGCGCGGCTCTCTACAGTACTTTTTGAA
GGTGAATGCTGTGATGACATACCGCTACCGCGAGCAGCGCGGCGCAGGGAACAGGATGAGCAGGGGATGATGCTGATGG
GGATTGCCAGGTAGAGCGTCAAGGGACGACCGTACGCGAATGATGCTGCGCGCGCTTGTGGTAGAGGTAGCTGGAC
TCCCGCGCGCGCGCAACTTGCAGAGTCCCGCGCACAGGCGCTCCACAGGCGGTGAACGCGCCATCACGAATAA
CGGATCGTCATGACAGGATCACAGGCGCAGCAGAAATGTCAGCACCGGTGACAGCGCGGCGAGGCGTATGTTCTGCA
GCGTGAGACACCGTGGGCGGTGAGATAGCGAAGCTCAGCGGGCTCCGTGGCGCGGCTGCGCGATGGTAGTCATGTTG
GTCATCCAGTCGACCATCCCGTCTTACGAACAACAGTCATAGGCCAACTGGGCGAGCAGGTGGCGGTTCCAGCTTC
CTCTGACGACGAGCTGTGTAGCAGCCCTGCGACAGCCAACTGAGTTCGACTCAAACATGGCTGGGCTGCTTCC
AGCCCGGTGAGGCGAGAAATAACAGGCATACCCACTCCAGTATGATGCGCGGATCAGCGACCCGAAACATCACGCGG
ATTCAGGAGCAGGATAGGGGGCAATCGAAATGCCAAGTGGTGGCGGTTTCTGCGGTAGCGCGGGGAGAAATG
AACAAAGATGTTCAAAGCTTCTATAGGAACTGTGTTGTGAACCTCCATAGAACTCCCTTTTACACGATGACGTTGTT
GCTTGGAGTGTGTATCCACTAGTTCTCGGGGAGGCGCTCACTGTAGATTTTTTGGGGGGTATCTCGGCATAGGGGGAT
TGAATAGCGCTGTGCTGCTGAGCTCTGAGACCTATTAGGAATCATCTGATAGAGGAATTTTTCCATGCCGTGAGTCTTT
CGGCATGAAGTTGAAGATTTCTACAGCTTTTGTAGCGGTTTCTCCACTCCGAGTCTGCGCTCAAGCTGCTCCGCA
TGCTCGGTGCGCGCTTGTGTTGTGCGTGTATGACGCTGATATTCAAGTGTGAGCATGTTTAAACCATCAGGTGCTCC
CTCAGTCGCGCAAGCTATGAATGTGGCTATGTACGAAGCGCAGCTTTATTTGAGCAGCGCAGGCGTGTCTCAATCACTT
GAGCGCAATGTGCTGCCCTTGGCCGCGGTGAGGCGCTCGTCAACGAAGCGCCGAACATGTGAGCATCTGCGCTGTA
GTGACGAGGGGCGAGGTCTGCTATTGACCGCTCGCAGCTCGGTGATCTCGGGGAAAGCGGCTGGCACTGATGATCTG
GTCGATACCGACAAAGGCGCTTGGTTTACCGGCTTACCGCGATGGTAGGCTCGGCGAGCATATCCAGCAGATAAC
CAAAGAGGTGTACCGAGCTTGTGCGGACTCCGTGCGCGCTGTCTACGCGGTGACTGACGTTGGTACCCCTCAACGGC
GTACCTTTTGAATCTTAGCGGATGAGCGGGCGAGGGGTGGTAGGCTGGAGATTCTCGCGGAAGACCTCGATTGCG
ATGTTGCGCGGAATGATGCGGAAACTACATGCTGCTGGAATCAGCATGGGCGAGTCTGACTCGCTACGGACGAGAGGC
GCTGGGAGCGGTGCGTGGCGACGCTTTTGGCTGGAGACGGCTTCGGTTTATCGGTGCTGGCCCACTGCCGAGCATA
TGGTGCTTTTCCAGCAGTGGGTCTTCCAGCTGGGATCTGATCTATCATACGCTATCGGTGCGCTGTTGCTGGCTCTG
TGGCTCCCTCTGTTACTTGCCTTGCCTTGCCTCGCAGTCCGATCTACTGCTATGGCTGGTGGGAGCATCGAGCG
ACGCTTGATAGAGCCGCAAGCGACGCTTGAAGCATGAAGGAGAGCGAAGCCTTTCCCGTGCAAGTTATCCAGGCGC
CGCCGCTCGCGCTGTGCGTGTGCGTGTGCGGACGCGCAGTGGTCTGGAATATCCCGAGGCGCGCAATGGCTGGGT
GATAGCGAGGCGATTGCCACGACGCGCGGAGATGGAATTTCCAGGCGTTCGAGGAGGTGTGAAGTGTCTGGAGAAGA
ACTGGAACCGAGGCGAGGCTACATCTCAATTACAGCCCAACCGCTATAACGGTGAAGACGATTTGTTCTGCGG
CCTTCAGTGAAATCAGTGACGCAAGCGGATGGAGGCGGAACTGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAATGAA
GCCAAGACGCTGTTTCTGCCACCATGAGCCATGAAATCCGACACCTCTGTACGGCATGCTTGGCAGCTTGAGCTGCT
TGCGGCTACCGAGCTGAGTCGCGCAGGCGGTTACCTAAAGGCAATCCAGCATCTCTGTCGACCTGCTGCAACTGA
TCAGCGATGCTTACGATATCCAAGATAGAGGCGGCAACTGGAACCTAGAGTGGTGAATTTCTCCCGCTGGAATTG
ACCGAAGAGGTGTCGAGTCTGTTACCGGTGCGCGCAGGCGGCTGCAAGTGTATACCTGCCCTCTCTGCGGAGCT
GCCGCTGCGCATGCGGGGGCGCGCGCTGATCCGCGAGATTCTCAACAACTGCTGAGCAACGCGGTGAAGTTCACCG
ACAATGGCTATGTCAACGTCCACCTGAAGGCCAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGAT
ACCGCATGGGATCAACGTGAGGATCAGCGCGCTGTTTGAAGCTTACAGGAGTGAATGAGTGGTGGTGGTGGTGGTGGT
CGCAGGCACGGGCTCGGCTTGTGATCAGCCAGCGCTGTTTGAAGCTTACAGGATACGCGGCTCCGAGCATCCGCT
TGGGTTGGGAGCAGCTTATGCTCAGGCTTCCGCTTGGCGGATCGCGATGACGCTGAGCCGAGGCTGAGCCGAGGCT
TGCGCGCTCAAGTGTGGCGCTGTCCGCGACCTAACGGAATGCTGTGGTGGTGGATCTCCCGTGGGTGGAAGGGC
CATGGTCGCGACGCGGAGTGTGCTGAGCAGGCGGACGCGACCTCGTGTGGTCAAAGTGTACTGCTGGAGGGGGCGC
CGATGTTCAAGCATGGCAGGATGCGGGGTGAGCTTTCCCTCAGGCTGATATGAGCGCGAGGCAAGGGCGCGAC

Figure 2

FIGURE 2C

TGGCTGCTCGGGCTCAACAACCTGAACGGCTGCATCGTGCTCTGGGCTGGCCATGGGCGTCTCGCTGATCCTTCGAC
GCCGCGATACGGCTGGCTCCGTTGCGCAATCTAGGTCTCCGCGTCTAGTGGTGGAGGATAACGCGATCAACAGTTGA
TCTTGAGGGACAGATGGAAGCGCTGGGCTGCAGCGTGGAGCTGCTCTTCGATGGTCCGAGGCGTGTCTGCACTGCCAG
ACGGCTGCTTTCGACGTGGTGTCTACCGATATCAACATGCCGAACATGAACGGATACGAGCTAACCGGGAGCTACGGCG
CCAAGGTTCCGGCAGCCGATCATCGGCGCAGCGTGAACGCCATGCGTGAGGAGCGCAGCGCTGCATGTCCGCGGGA
TGAACGATTGCTGCTCAAAACCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCTCAAGGTGGATCGATGA
GCTGGAATCCTATCGGGTGTGGTGGTCAAGATCAGCCGTTTACGCGGAATACCTGCTCAACCTGTTTTCGCGAGCGC
GGCGTGCACTACCTGGTAGGTGCCGCGACGGCGCGAGGCGTTGCGCTGCTGAAGCAGGACAGGTTCGACCTGATCTCT
CAGCGATCTGATGATGCCGGGATGGATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCTGCCGAAGCTGG
CGCTGATGAGCTCTCTGTCGACGCGGATGATGCTCAGTGCCAGCGGGTCCGCCAGAGTCTCGGCTTGTCCGTTAATCGAC
CTTGTGCCAAGCCGACTCTGCCAAGGCCATCGGCCAATCTTGGAAACCTGGAAAGATGCTCAGGACGAAGCTGGA
CGCGAAACCGACGAGACTCCGATGGGCGCAGCGGCTTGGATGCGCTGCATAACGAGCAACTGGTGACCTGGTTC
AGGCTAAGAAATCCCTCCACACCGGGCGCATAGTCGGCGCGAGGCGTTGATACGCTGGAGCCACCGCAGCATGGCCTG
TTGCTGCCAGCTGTTTCATGAGTGATGTCAGCGCTACCGGTCTGCACGAGGCGTTGCTCTGGCGCGTGTCTGAACAGAC
CCTGAACGCCCAGGAATCTGGCGCAGGGCGGTTACGAGATTCCGGTTTCGGTGAATCTGCCGCGCACCTGCTCGATA
ACCAGGAATCTCCGATCGACTCTATGAGTACGTCGGCGCTCGCGGGCTGTACCAGCTCACTATGTTTCGAGTTGAC
GAGAGCAGTGTCAAACTCTGTCAAGTAACACTATGCAGGTGCTGTCTGCTTGGCGATGAAGGGTTTCGGATTGGCCCA
GGACGACTTTGGCCAGGGTTACAGCTCGTTCTATAACCTGGTCACGACGCTTTACCGAGCTGAAGATCGACCGCTCCC
TAGTCCAGGATGCGTAGAGGATAACGGCTCAATGCAGCTGTATCAGTTGTATTGAGTTGGGTCAACCGCTGAATCTC
GAGTGGTGGCGAAGCGTGGAGACTTCGAGGAACCTGAATCTTCTTCGTCGCTTGGCTGCGACCGGGCGAGGTTT
CCTGATTTCTAAGCAGTGTCTGCTCGTGAAGTTCGAGCGGCTTAAGGAGGACCGGCCCCAGCCTCCTTGTAAACGCA
GTATCCCATATATCGCGGAGTGCATCGCAGCCAAACCGCTCAGCGCAACAGTGTCTGTAAGGAGGCGCATCCCGTGA
AGTCTGCTAGTGCCTTGGAGCAGCACAACAACTTTTCTCAATGGACAACCTCTCGCAGAGCTGAGCATCGGCTTG
ATCTGTGTGGTGGTGTGACCGTATTGCTGTTTACGATCTGTTACTGGTCTGGGAGATTGTTTTCAGGAGGAGGAGGA
CAAAGTCTCCTTCCACTTACCGGTATGATGGATGTTATACGGGAGCATGAGGTATTTCTTGGGCGCATCGCTCGAAAAA
CGCACAAGACCACCCAGAAGTACGACTATGAGTGGTGCCTTTCAGCGGCACTTGTGGCAAAGGAAAAAGGATTAGCG
GTCTATGAGGGACGGGAGTTTCTTTGCTATGCCATTTCTACTGGCTACCAAGCAGCGGTTGAGCGCGGATTCTCGGG
AGATCCGTTTTCGCTCGGTGATTGCTCGCAATTTCTACGGAAGCTTTCGAGTGTTCGCGCTATCCCGCGCCACAGT
TACTGATCTTTGATCTTTCGCGCAGCACCGCCTGGCAGTGCCTGATTCCTCCACAGCGCAGCTGACAGGTGAGC
GGAAGCTATCCGATGATAGTCGAGCGCATCTGGCGCGCTTGGCAGCCCGCGGTGGGGAGGAGCTCAGCGTGTCCA
TTGGATACCGCTGATCGCTATCGCACTCGCGCTGGAGATGTTGGGAGTGCCTCGCGGTTGATCTGCCGGAACACTCT
GGTGGCAGCAGAGCCGAACATCTGATCATCGCTGCGAGCTGCTTGTATCTCAGGCGAATCAATGACTTCGAACAGTTG
GTTGAGCGCGCGCATTCGATTTCGTACAGCTGGTATCGCCGATGGCAGGATGTTGCTCGCGCGCGCCCTGCGACCGG
CCTGAGGGATGGCTGAACCTCACCCGACAGGGGTGCGCGTTCACTGCTCAGCCAGCTGAGAACGGTTGGTTCGCGG
TCTACGAAACCGACTACGGCAATTTCTTTCGCACTCCCGGTGGTGGCAGGTCTGCTGCTGACCCCGCGCTGCTC
CTGGCGGTTGGCTCGGGATGCGTTGGTACACAGCAGCGTCTCAACCGGTGCATCGGGCGCACCGGCACTGGTGA
GAGCGACACCTTCAGCCGAGCGTGTATACAGACCGCGCGGTGGTCTGTTGGTGTGCTGACCCAGGATGACAGCAACTGG
TGACTGCAACCACTTGGCGCGCCAGTGGCTGGCGGGCCACGGAGATCTTGGGCTGACTTCAACTGGAAGCTTTTC
GATGCGCGTGGGAGGTACAGGAGACATCTGTATCAGGTGCTGGGCGCTATTTCGAGACCGCTTTCGCGGCGACCGG
CTATGCCGCGCAGCAGGCGGTACTGTGCGTATTCAACGACATCACGGTCCACTGCGAGGCGGAGACCGCGCTGTCCAATG
CGAAGCGAGCAGCGGATGCCGCCAGCCAGGCCAAGACCTGTTCTTGGCCCGCATGAGCCATGAAATCCGTAATCCCTG
TACGGTGTCTTGGCACCCCTGGAGTTGCTCGACCTGACCACTGACAGCGGCAACCGCGCTACCTACGACCATCCA
GAGTTCGTCTGCGACGCTCATGCAACTGATTAGCGATGTGCTGGATGTCTCGAAGATCGAAGCGGGCAGATGGCTCTGA
CCCTGGCGCGCTTCAATCCGCTGGACCTAGTGCGGGAGTGTCTGGCAACTTTCGCGCAGCGCCATGGCCAAGGACCTG
CAGTTCATATGCTGCATCGACACCGAAGTCCGGCGCAACTGATCGGTGACGTGACCGGATTTCGCGAGGTGCTCAATAA
CTTGGTGAATAACCGCTGAAAGTTTACCGATATCGGACGGGTGGTCTGCGCGTGAAGTTGCTCTCCGCAATGATGGTC
GAGCCCTGTTGCAAGTGGCAGGTGCGCGACACCGGTATCGGTATCGCACAGCAACAGCAGGAGCGCTTGTTCGAGGCGTTC
TACAGGTTTCGGGAGCGCACCATGCCGCGCGCACGGGCTAGGACTGTGATCTGCTGGCATCTGGCGGAAATGATGGG
CGGTACCTGCGAATGGTCAAGCAGAGAGGCTCGGCGAGCTTACGCTGCTGCTGAGTTGCCGAGGACGAACAGT
CCGGCTGGCTTGGCGCGGGCTCTTGAATCCGCTTGGCTCCATGTGCGCTCGCCCGTGGGAGCTAGCCGAGCAGC
GTAGGGGCTGGCTGAAAGCTTGGGCTGCAAGGTCAAGCAGCGCGAGGCGGCGCCCTCCGAGCTGGAGACTTGTGTCT
TCTGGAGTGTGCTGCGATGGCGCGCGGCTGCTTCTTCCGCTGGCCAGGCGCCCGGTGCGCGCTCCATGGATGCGC
CTTGGCAGCGGAGCTGCGTGAGGACGGCTGGCGTGTGCGCTGCACAACCTGGCGGGAATCGGCCAGGCGCTGGCGCAG
GCTCTGGGTGGCGATATCCCGAGCAACCGCGGCAATGCTGCGCGCGCTCGGGGAGACTCGAAGTGAAGTGTGGT
CGCCGAGGACAACCCAGTCAACAGGCGCTGCTTCCGAGCAACTGGAAGAGCTGGGTGTGCGGTGAGCTTTCGCGCG
ATGGGCGGCGAGCCCTGCAGCTGTTTCGACAGTGGTCTGCTTGCAGCTCTGCTCAGCGAGCTCAACATGCCGAACATGACC
GGCTACGAAGTGAACAGGCGCTGCGCGAAGGCGAGAGCTGCGGATCATCGGCGTGACCGCCAACGCCCTGCGGAG
AGAGGGCGAGCGCTGCCGGGAGTGGGAATGAACAGTTGGCTGGTGAAGCCGATCACTCTGCATACCTTGCATGAATGC
TCAGTGAGTTGCTGCGCAGGTGTGCTGCTTCCGCGCAAGCGCGAGACTTGGCCCGCGCGCAGCTCGACGACGCT
CTCTACCGCAGGTGCCGGAACGATGCGCGCTTTCTTTCGAGACATGGGCAAGGACCTGGAGGCGCGCGGCAAGC
GATTTCGCGCAACGACCCGAAGGGCTGCGAGCAGGCTTCCAGCTGATGGCGGCTCCCTGGCGGTGATGCTGCGCGAA
CGCTGGTGGTGTGTCAGGGCGCGGAGGAGCTGCTGGAGTTCGCGCTTGAATGTTTCGCGGTGGAGATTGGCGAG
GTGCTCGTTTATATCGAGCAGGCGTGGAGTTGTGAGAAAGACGGGCTGATGCTGCGGGGTCAATAGTTGGAATTAGA
ACACAAGAGAAGCTATGAGTAAGCTCAAGATAGTACTGGCCGATGACCATCCGATCTGCTGATGGGCGTATGCGACAT
GCTCGAGCGCGAGCTCGGTTTCGAGGTGGTGGGCGAGGCTTCCAGCGGCAAGCTGGTTCGAGGTGTCGCGGACGAGCG
AGCGCATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTACGGCGATGGAAGTGAAGTGAAGCGGCTTCGCGCGCTG
TTGCGCAACTTCTCTGTAATAGGTGCTCATCTTACCATGGTGGCAACCGCTGATCTCTGACAGCTCTACGATCA
CGGGGTGTCGCGGTGGTGTGAAGAGCGGCAACTCGACGAGTGTCTTGGCGCTCGAGTGGTGAAGCAGAACCGCG
TCTACCGGGCGCGCAACTGTCTGACCCGACAGTGTCTTGGCGAACCAGCAGGAAGTGAAGCGGCTTCGCGCGCTT
TCGATGAAGGAGTTCGAAGTACTCCGTCACTTCTTTCGCGCAGCAACGTCTGCGATATCGACGCGCTGCTGAACGTA
CGTCAAGACCGTAAGCAGCAGAAAGTCTCGCGCATGCGCAAGCTGGAAGTGAACAGCGACAGGCTTGTGACCTTCT

FIGURE 2D

CGGTGCATGCCAACTTGTTCCATTGACAAAGTTTCCACCGCGGAGCGTTTCGAGACGCGCGCGGATGTCTCAGTGAAGCA
GGCTCCGTCCTGCTGGAGATTGCCGCGCTCATCGATCCAGCTGAAGCGCACTTCGGCTTGGCCCGCGGAGCGCTTGG
CGGAGCGCCAGGACGCGCTTGGAGAAGGGGGCAACAGTTTCGGCCGCGCCAGGCGGGTGTGGCCGCAAGCTTCGGCGCC
GGCGAGGCTGATGAAATAGGGAGTAGGGTTCTCCATCTGTAGTGTTCGCCCGCGCCGCTCGAACCTGAGCTGGTGTGGCG
CTTCTTCGCGACCTCTGGCCAGAGGGCGAGGGCGGTAGATGACTTTCATCTGGGTGCGCAGTGAACCGTCAGCCGTGAC
TGTCTTCGCGCAGCAGCCCGGTGGCTTTCGCGCGGATTCGTAGAGGTTCAACAGTACAGCGATTTCGGGTCCCTTGG
CAGCGCTTGGCCGGTCCGAGCAGGCGCAGGCTGCGTGTCTGGGTTCGAGGCGGAACACCGCGGTAGCGGCATGA
TCCGCGCGCGGGCAGACTCGCGTGTTCGAGTTTCGGGGCGCCATCGTCGATCCAGGTCTGCACTATGACCGGATAGGCATTC
TGGTTGACCAGCAGCAGCGACGTCTCGCGGTGCCCTTCGGCGATTATTGCCGAGTGGCTCGCGGTGACTTCGGCTCG
GGCTCCAGCAGCGGAGCATTGCTAGGCCGAGCAGTGCCTTCGAGCAGGATCTTACTGGACACGAATCACTACCTGC
GCGTGGCTCCACCCGTTCGGGGTAACGCTGGGCATGCTGCCAACGGTGAAGTTTTCAGGCGCGCACGGAAGGTCTC
GCTGTAGTGGCTGATGCTTCCGCTTCGCGAGTCTTGTTCGAGGCGTTCGCGATCACTGGATACAGCCCGTGTTCG
CATTGCTGCCCATCGCGAATCTCGCTTACCAGCAGGTTTCATCGACTGCCGTTCGCGCAGCAAGCGGATACCCAGCCT
TGGGCCATACAGGCTGGCGGTAGCGGTTCGGAAGTAGGTAGGAGAGCCCGCATTCGCGTTCCTCAGGCCAGTGCAGC
GGACCGGCTTCGCGCCCGCTGAAGCCCTGATACCCAGTGCAGTACCGCTGGAGGCGGTGCTGTTGATCACTCCGTTT
GGCACTTGAAGGCCCATCGAAGGCGATCTCACGTTTCGCGCGGCAATTAATCTCACTAGGGAATTCAGCGGAACACA
ACGTGGGCGGTGACGTTGGTTACGAACAGGTGGGATAGCGCTTGGAGCTCACGTGCTGTGTACAGGCCGATGGCGCCG
CCAGTTGTAGTCCAGCCGAGGTAGTTGGCGTTATGGCTGGCGCCGACGTTGGGGTAGGCCAGGCCGGACCTTGTATGG
CGATGATGACAGCGGCTGGGTGTAGGCGTACAACCGCAGGCGGTAGTTCGGCGAGTAGTAGCGGATGAATCCAGCGGA
GCGCGGAACAGTTCGCGCGTACCGCGCTGAGGTTCTTCGCTTGAAGCTGAAAGCTTCGCGAGTCGATATCGAGCC
GCCAGCAGACGCTTCGCGCCAGACATCGGTGAATACTGCCCGTTTCACGTTGGGTAGGCGCAGCAGCAGTTCGGGCC
AGGCGGTGCGATAGGCGGACTGCAGGCCAATCGAGTTGCCCGCACTGTTCGCTTCTAGGTACAGCCGCTGTAGAGATCG
TCGGCATTGGTGGAGAACATCTCGTACACATCGTCTTCGAGCAGCAGCGGAAGAACGCGCTCTGGGTGATAGCCGCG
GCTGTTCGCGTATTTGCTGAGCGCGCCACCGAGCTGGCATTCAAGCTACCGTTCGGGTGGAAGCTGGCGCTGCTGACGT
TGATCACGCTCGGTAGACCGAGGAACCGTTGCAGGTATCGCAGGCTTCGGCCAGGTGCCGCGAGGCTTCACCGGA
CGTATCTGGGTGTGGAGGTAGTCTGTGGCGTTGCCCTACCGCGTACCTTGTAGCAGGTTCGCCAGCTAGACGAGCGAG
CAGCGCGAGGAGCAGAGGCGAGGACGCGCAGGATGTTTTATGGTGCCTTCGAGACGGGCGTGCAGGTGCCCTGCAGGCG
GATCACCGCTTTCGCGGTTTCGATAGCTTGGCAGGCGGAGTGTTCAGGTCGTAATCCAGCAGACAGGCGCTGCTGACGT
GTTGCTCCATTGACGCGCAGGCGGCCCTTGTTCGCTTCGGCGCGGCGATAGACCTGGCCGCTTGACCGACCATGCCG
ATGTTACACCTTGGCTGTTCGAGCAGATTGGCCCTTACGCGTAGCGCGTTCGCTTCGCGAGTTGGCGCTGGATTAGCAA
TGGGTGACCGGTGAGTGTGCGAACTTACATGCACGCGCGCGCGCTATGGCGCGATCTTCGCTCGGTCTCCAGCA
GCTCGGCTCTTCGCGCATGCCCTGCGGATCGAGGTGACCGGTTGTAGCGGTAGGCGGAGAGTGTAGGCAACCGCG
TAGCCATTGCGGTTGACGCGCGCGCTTACACCGCGTAGCCAGCTTCGCTGGCGCTTGGCTTCGACAGGCGGAA
AGTGTTCGCGCAGTAGGGGCGAGGCTCAGGCGCGCGCATGCAGCAGCGCGCGCGCGCAGGCGGTGTGAGCTGCC
GCTAGCTGCTGTTCTCGGCATAGCCGCGTTCAGCGTCGCCACCGAGGTCTGTTCTGCGAGTTCGCTTGAATCGCTG
CCGTGGTTGTGCTGTTCGCGCGCGCGGCAATCGCTAGCTCAGGCTGCGGTCTTCGTCGAGGCTGCGGTGAGGCGCT
CTGGTAGCTGTTCGCGCGGCTGTTCGCGGAACGCGCGGAGGCCATCGCGTTCGAGATTTCGGGCTGCTGGAGGACCCCA
CGGCGATGGAAGTGTAGGTTAAGGTGTGCTACGTTTCGCTGCGGTGCGCGAGGTTGTATTGCGCGCGGAGTACGCGG
TCGTAGTCTGCTTCGAGATCGCGGTACAGGTGGTCTGCTGGCGGAATAGGCGAGGTTGTAGCTGAGCTGGCGCCAGT
GTTGCTGTAGCGAACTGCAACTGGGTGTTCGCGCTTTCGCGTGTAGTGTGCTGGTTCGGACAGATACAGGT
TGCGGTAGCCCCCAGGCGCTGGTTGACAGCAGGTTGAATGTTGCGTGTGTTAGCTGCTGGAGTTCAGGAGTTCG
TTGTGCTCATCGCGTGGCGCGCGGAAGCGCGTTCGCGAGGTTCGCGATAGCCCTCGGTGGAATAGCGGTAGCGAGCCAG
GGTGAGAGTGGTCTGGGTGCGCTGGAAGTTCGCGTGTAGTTTCAGACGACACGCGAGCCCTGCTTCGCGCTGGCCTCT
CCACCGTGGCATGGGAAAGATGCTGTTGAAGCCGAAGGCTCCGTAGGGCGTGGCGAGCAGCCCTCCGCGAGCAGCGCC
AGATAGTCTTCGCGCAGGCGCGAGCCGAGGTTGGCGGTTAGCGAGTTGGTCAGGCGCGCTGATAGGTGAAGTTCGCCGA
CAGGTCTGTTGCTGCTGCGCATACTGCGCGCTTGGCGAGGTGGCGCTGTAGCGCGAGATGCCCGGCGCATGGATTGCG
GAACCGCGAAAGGGCAGCGAAAGCGCGAGACGCTGCCGTTCGCGCTCAATCACTTGGACATCCAGATCCCGCTGTAG
GCGGTGCGCATAGAGTCTGTTGATGACGAAGGACCGGGAGCGACGCGGATTCTGTGACCTTCTTGGCGTTCTGGCTGAT
GACCACCGTTCGCTGTTGCTGTACCGGTACTGTGGAGCATAGCGCGGTTTCGAGTCTGCGAGCATGCGGTGCT
CGCTGGCCAGGCGCAGCGCGGATAAACCAGGCTGCCGAACAGATTCGCTCGGTGAACTCTCGCCAGGCTCAGTTTCG
CTGCGCAGGCTGGCAGCGCGCTGGGCATAGGTGCGGATGCTGTTCCAGTTCGCTGCGGGTGTTCGCGCATAGCTGGA
GTAGCTGTAGTTGAGTGGTGGCGCAGGCGCAAGGCCGAATGATGCCCGCTCAGCCCCAGGTAGCCATAGTCCG
AGTCCCGCTTGGCGCTACTCCGTGAAGCTCGAGCGATAGAAGTGGCGTTCGAGTTGACGAAGCCATGCTACTGCCA
GCGTCCCACTGCTCGGAGCCACGATAGCCGCGTGGCTTGTATTCAGCAGGCGCTGGGGCACCGAGAGATCGAGGCGCAG
AAGGGCGGTGTTCGAGATTGAAGGTTCGAGCTGGCAGGCGTTCTTCAGCGCCAGGCAATCTCCAGGCGCTTCGACGCGG
CTCAGAGGCGGGGCGGACGCGCAGCGTTCGCGCGCAAGAAGCGTTTCGCGAAGCAGGTTTACGCGCGCGCGCGCGGGG
CGGAATCGATCTGTTGTCGACTGGCGTAGCTGCCGTTGAGATAGACATCGACATGTTAGGTTTCGGGGTCCACCTGGTT
GGCCCGTTGAAAGCTTCTAGGTTCCGCGCGGCGAGGCCCATCAACAGGCTGTCTGCGAAACGATCTGCGCCT
CTCCATTGGCCAGGCGCGCAAGGCCAGCGCTAGGCCCGCCAGAGCCAGCGGAGTTTGGCGGTGCGCGCTGCTCAACG
TGGCAGAACATGCTGGACCTTCATTTCGCGCGCGTGTAGTCATTGATCAAGAGGGCGTTCAACTGGACCGGTCTGGGGCGA
GCGCGAAGGGCGTTCCGCGCCCACTCGGCTGGCCTTTGGGCGGGACCATGCTCGACAGCAGCGCCAGCGAGCGTTTCG
CCGATGCTCAGCATCGCGCGCGGAGGCGGTAGTGGCGTTCGCGCATCCCGCACAGCGGATGTCGCGCTCGC
CTGCACAGGCTGAAGTGAAGTGTCTGACTAGCTGCTCCGAGCGCCAAAGCAGCGCGGGCGGCGATAGAACAGCTTCA
GTCGACTGCGCTGATCAGCAGCATCTGGTTGCGCTTCGCGCTTGGCCAGGTTCGCGGGCGGGATCTGCGAGCATTGAGA
TGAACAAAGATTTCGCGATCTTCGCGCAGCGGAGCGCGGTATACAGGAGGCGCAGGCTGTCGCGCTGCGGGGGCTAT
GCGCGTACCGCTGGGCTGACAGAAAGGGCGCGTTGCACTTCGCGGTGGAGGAGGGTTCGCTGTTGCTGCTCAGG
CTTGATCAGGTTGGGAATGCGTCTGATTGATCATCTGCACGCTTTTTCCGGGCGTTCGCGGGATAAATCAGCGGA
GTACCGATGATCACCAGCTGCTTTGAGCGGGATGGCGGTATCAGCAGGCGAGCGCCAGCGTAGCCGACGAGCGG
AGGAGTGGAAAGTATTACATCTGTTACCTTGGCGCGTGGCGCTCAGGAGGGGCTCGGTTAGTTCGCGCGGAACCTGG
GAGGGCGCGCTGGGCGGCGCTCCGCTATTACTCGTAAGCAGGCTGAGGTACCGAGCGAGCACCGCTCCGAGT
GACGGTGGTTCGCTTCGAGATGTAGCGACGCGATAGTCTGAGTGGTGTGCTGCTTCGCGCTTCGCGAGCAGATGTCAC

FIGURE 2E

CGGCAGCCACGGCCGACCCGCGGCCAGATTGACCTCGCTGCCGCCCGGTGCCGTGAGTAGCTGACGCGTACGCCGGT
GGCCGCGGCGGACCGCGTATTGGTCAGATTGCCGCGCGTGGTCGGATTGACCGCCTGGAAACATCACGCTGAAGTGCTCCTC
GGCATCGGCGCGCGGAGCGGCAACCGGTGAGTTCAGGGTGAAGCTGGTTTCCCCGCGGGTTTGCCGACTGCGCCGT
CAAGAGCGCTTACCGGTACCGGTGTCAGTATCACGGTCGGGTGCGTTCGTCGCGACGACGCGGAAACAGGCTTCGGTCG
GTCACCTTCTCCGTGGAAGGTGATGGTATTGGCAGCGCGCGGATACCCGGCAATGCGAGAAATGCGGCGATAGTCAGGGC
AGTACTTTGAGGTTTCATTTCAGGGTTCCTCTGTTGAGTTTTTCTCCGAGCAATGGTCTGTAAAGAGCTATCCGGGC
GTATTTAGGACGTTTCTTAAAGGACGAAAGTGGGTTTCTCGCTCTCGGCAATGTTTTTACTGCAATGGACTACCCCTG
CTCCAATAGACATTTCCGGGCTCATAGCGAGGCGCTTTCAAAGGCTCGGACTGACGCCACCCAGATGACTGTGTCGCC
CGGTTTCGGGTGTAGAAGACCTCGATGTAGTCAAAACATCCGCCCGGCCATGTCTCGGGTTTTGTAGATGCGTTTTGCG
GATACGCTCTTTCTTAAACTACTGAAGAAGGATTCCGCCACGGCATATCCCAACAATGCGCGCGCGGCTCATGTGCG
CTCCAGGTTGTTGGCAGCACAGAAGCGCTTCAGTCGTCACCTACCGTGGGACTCCTATAGGTACTTTTTGAAGGTGGAT
GCTGTGATGCACACCGTACGCCGAGCAGTGGCGCGGCGGCAACAGGATGAGCAGGGGATTGATGTTGATGGGAATTGC
CAGGTAGAGCGTCCAGGGACGACCGTAGCGGAATGATGCTGCCGCGCGCTTGTGGTAGAGGTAGCTGGACTCCCGGC
CGGCGCGCAACTTGGCAGGTCGCCGCGCACCGGCGCTCCACGAGCGGCTGAACGCGGCGCATCAGCAATACCGGATC
GTCATGACCGGATCACGAGCGCACGAGCAATGTGACGACCGGTGACAGCGCGGCCAGGCGGTAGTTCTGACGCGTGA
GACACCTGGGCGGTGAGATAGCAACGTCAGCGGGCTCCGTCGGCGCGCTGCGCGATGGTAGTCATGTTGATGGGAATTGC
AGTCGACCATCCCGGTCTTCAGCAACACAGTCATAGGCCAACTGGGCCAGCCAGGTGGCGGTTGACCTGGCTCCTGC
ACGACGACGCTGTGTAGCAGCCCTGCGACAGCCAACTGAGTTCGTAATCAACATGGCTGGGCGATGCTTCCAGCCCGC
GTCAGGCGCAGAAGAAATACAGCAAACTCCAGTATGATTGCGCGGATCAGCGACCCGAACATCAGCCGAGAAATT
GGAAGGCGTACTGATGATGGTGGCGAGAAGGTTCTTGGCTCTCTGTTGCTGCTCTGCTCGTTGAGTAGTTCAGCC
ATCAGGCGCGCATCTCAGCGGTGCCGAGGTGACGGAATCGTCGACAGGTCCTGGGCGAGACCATCTGCGAGGCGGAG
TATCCCGCGCCTCCCACTCGCTGTTGGCGGCTTGGCCCTGCGCATACCGCGAGCCAGCTCCTGCGAGGCTTTTCGG
CATACCTCGTCGGGGGCGACCGCGCGGCGGCTTCGAGATCTTCAGAGATTGCCGCTCGAGTAGCGCGAACGCTT
GTCTTTGGGCGAGCAATGATGTGGGCGGCTCGAGCATCGGCACGCTGGTCATCTGCACTGGTCGTCGCGAGCTGGAG
GTGAAGGTACCTTCTGTTGTTGTTGATCGCGTGGTGGCGCGCTCGCGCGCGTGGTGGTAGATCTGGACCTTGGG
GAGCTGATTGGTAAGGAGTTCGGCGGTGGCGGTCTCGCGCACCGCGAGCATGAACAGGTGTTGAAGTTGCCGATGATCT
GACCGGCTTCGCGCGGAGCGCATCTTGGCTCGATGTGCTCATGGTCTGGGTGATGGCGTACCTGCAACGCGCGGCG
CCGCGCGCTTGTGACCATGGGGATGAATCGTCGCAATCAGCTCGTTGAATCGTCGCGATGCGAGTTGATGCGGAC
CTTGGCGCTGGCGAGCGAGCCGGGCGAGGCGCTCATCGACCATGCTTGTAGATGTGACCCGCTACCGAGACAGGTGCG
TGAACATGGAGTTGCCACCGCGCGGCGACCTCGGTATCCGATAGTGGTTCGAGGCGGACGTAGACCAAGCGCGCTTG
CGGATGAGCTCATCCAGTCGAAGATCGGCGCGGATCGTTGAGGTCGCGTAGTTGGGCGAAAGAACTCCGAGATCCG
CCCGGTAGTGAGTTTCTCAGTAGCGGCGAGCGAGGCGGACGATCTTGTGCAAGTAGGTCTTGTGTCGAGCGCACCGCG
TCTTCAAGCTTCCATGACCGGCTCGCGATGCGTTTCTGTGTCAGGTACTGGTCGATGGCTACGACCCGCGAGGGCGCT
CCTTTCTATGTTGAACGGGATGTTCTTGTGCTGAGCTTGGCTCGATCTGGATGATGGTGTCCAGGCGCTGGGATCGTG
CTCGTGTGATGATTTCTGCGCATATTCGACGAACACGATCGATGTTACGACGTCGCGGAGGATCTGCTGGTAGTCAG
GCCGATACCCAGCGCTGCGAGCGCGCGGCGATGATGTTGACGAACCGCGAGGCGAACTCGCGGAACCGCGCGCTGTTG
CCCTCGCGGAGAGCTGGCGCGCGAGCGCGGTGGCGACTCGGAGATCCGACCGAACCGGCGAGCGCGCTGTAGCGTGC
CGACAGGTGAGGATGACCGAGGTGGAACAGTAGAATCGTCCAGGCGCGCGGCGGCTTCGATTCACGTCATACGCT
TCAGCAGGTGAGCGTGGCTTTCGGGTGGAAGAGATCACCACTCGTAGTCCGCGTGTCTCTGCGCGCGCGGCGCGA
TAGCGGTGTGAACCGTCTGGGTCCGCGGCGCATCTTACCGCGCGCGCTCGTACCCGCGAGTGAGTGGCGGAATATC
CTGGGTGATGAACAGCTCCGCGAGGCGCGCTTACCAACCGCGGTGGTGGCGAGTACAGTGTGGCGGAGCGCGCTCG
CCAGTTGACGCGCTACGCTCTGTTTCGCGCGCTCGATGCCATGGAGCGGAGCAACCGCCACCGCGCGGCGCGGCGCT
GCGGGTTCCAGGCCACGTCCAGGCGAGTGGCTTGGCGACGAGCTTACGGGGGAGGGGCGAACTCGAGCTGCTTCTC
CAACCGCGCGCGCGCTCGTAGAGGGCGAGGGCTCGACGTAAAGGCGCAACTGGGGCAGGTAGGTATCTGCCAGGCGCT
GCGTATGCTTCTGCTCCAGCGAAATCCTTACCGATGAACAGGTGTTGTTGCTGACCGGCTCTCGGCGCTGGTCATC
GTGTAATTCGCGCTCGCGAATGTTCCGCGGTGAGCGGAGCACACGCGCGGCTGCTTACGCGCGACGATACCCAGCA
GGCGAAGCACAGCGCGGCCAGATGCCGAACAGCGGAGTGAGGAGAACGTCACGCGCGGACGATGCGAGCGCGGCTG
CGGTGAAGCACAGGCTGGTGGTGTAGAGCTTCCAGCGAGGCGGCAAGAGCGCTTCAACGGGTACTGGCCAGCCATGGCT
TACACTTCAGGTTGCGCGGCGCTGAGGCGGCTTGGCATGAAGTTCGAGCACCTGTTTCTGGCTGCTTACCTC
ATGATCAGGCGCATCTGCACAGCAGCTCGCAGAGGCTGACTCGTCTCAAGGCTTGGACGTACCTCACCGCGCT
AGCGTGAGATCAGTCCGAGCAACTCTGTGTACATGAGTTGGCTGCGCTGAGCGTGGCTCGGCTGCTTGTGACTAGGGTG
ATCAGATGAGTAGTCATGAGGACTCTACTGTGATGGCGGTGATGTGATGAGGGCGGGGTAATGCTGACGCGCTAGC
CGGTGACAGGTCGTCCGCGCGCGCGCAATATCTGAAGTCTTCCCCAGGCGCGGATTTCCGTCAGGCGCGCTT
GCTGGCCAGTTTCACTGCCAGGCGCACGGCTTGGAGGTCTGAAGCTCGTCGCTCGCTCTTTACGCGAGGTTTCGAGACA
GCGTGTGCTGCGCGACAGGAACAGCAGTTGACGGCTGGAGCGTTGATGGCGCGCCCTGGACCCGCTCTGGCGATAGG
CGGGCGGAGCGAACTGGAACGCACTGAGCCACGAGCGCCAGGTGCGCAGGTCCTGGTGGGAAGCGGATGGCTGCGG
ATCCAGGCTTGTAGTAGGGGAGCGCGGAGGCGCGCGGAGGCTTCGACAACGATCAGCCGCGCGGCGCGGCGGCTT
TCGCGAACGCGCTGCAAAATGCCAGCGTGGCAAGGAATCTTTCATGGAATTGGTACCAAGTAGGTTGACGCGGAGAACT
GGCTAAGGTGGCGGGAACCGTCTACGGTATTGGCGGCGAGGCGCGCGCGCGGCGGGCGGTGGTAGCGGCGGCAACT
TGGATCCAGTCCACCGTCTTCCGGGGCGGATGCATACAGGGCGTCCGCTGCTCGATCAGGATGTCGAGGTTGGCGTC
GAGGCTTCTGTAGCGGATCCAGGACTCGCAGGGCTGGAGAAACGATGTCGTTCCAGCGGATGTTGACCTGGCGAGGC
CGGAATCGACGCTTGGGCGCGTAGAGGTTGATCGCGCGAGTAGCGTGTGACGCGCGGCGGTGCGGTGCGGTAGTAG
TAAGATTTCGCGCGAGCTTCAATGTCCAGGGCGAGGAACGATTCGCGCGCGAGCAGGACCTTGTCTCGGTCAGCGC
GACCGAGTAGAGCACTCGGCGGGAACCCCTTGGGCGAGGCGGATTTGCTTGTACGCGCGGCGGTGATTCACGCGCGCT
GAGGCTGTTGACAGGACAGCAGGGTGGCCAACTGGAGGCTCGAACTACAGAGCTTGCATCGGCGCTTACCTCTCT
GTACAGGCGGGAAGGCTCGGCGCAGTCTAGGGCCATCCAGCGGCGCTCGTCATGATTGAGCGTGATCTGCTTGTG
CGAACCTTCTTGGGTCGATGCGCGCGAGGATTGCCAGCGCGCACTCGCTCTGCGTGGTCTGGTACCGACGAAGTA
GAGGTGGAATCCTTTCTGATGTTGACAGGTGCGGACCGCGTGGATGACGCGGTCAGTTGTCTTGGACGAACAGG
CCAGGCGCGCGTCTGTCGAACGCTGGGCTCATGTTGGCGTGGCGCGCGGTTGGTGTGCTGCTTTCGTTGAGGCGG
ATCACCCCTCGCTGGATAGGCGCGGCGGAAGGCTTCGTGATGTCGCGCTGGTAGGCGAGTTCTTCTCGACCCGCGG
CGGTCGCGCTGGACCTGTAGATCGGCATACCGCGCGGCTTCTGCGCATCGCGCTCGATGCCAGCGCGGTACGCG

FIGURE 2F

GATCAATACCAGGCGAGTAAGCGCCCCGGGGCCCTTGATGAGCGTCTGGTAGCGGGTCCATTCTGTGGTTGAGTCCC
CACTCCTGGGCCAGGTGCTCTGAGATGTCTGTTGTTTTCAGCGTGCCGAGTTGGCTACCGGACAGTTGGCTGCTGGCCCTC
TCGGCTTGGCCGCCCTCTTGGCCGGCTGCGGGATAGCTCAGCAGCAGCATCAGGCCGACAGGAGCGATACGGTTCTGA
TCATGGGCGGGGTCTATGGTTGGATGCGAACGCTGCGCGAGGTGCCGGCAGCTCGAAGTGGCGGTACCATCGTCAAG
GTCGGTCAGTCGCCAGGTGCTGCCGGCAGCGGCATCTCCCGCGCAATGAGGTAGATCTGGCTGAGCTGGGTGGATCCCG
GAGGTGCAACCGACAGAAACCGTTCTCCCGCGCACTACTCCAGCCCAAGGATCGAGAAGGGCGGGGGCTCCATCGGTTTG
GGTTTCGGCTTGGGCTTGGCGGGTATGGGGCGTTTGGCGCTGGCAGCATCAGCGCGCGCCCTTGGCTGCTTTCTTG
AAGCGTGCGGACAGAACCGTCCAGTGTCTCCACGGTTGCTTGGAGCACCAGAGGTGCCGGCGAGCGGTGGTCTGAG
CCAGGTTCTCGACGGCATCGGAGGCTGCTTGGCGAACGCCCTGCGCAGCGTCAATTGCGTTGGACAGCGCTGCTGGCCT
GAACGGAAGTCTCTGTTGCTGACCAAGGTGCTGCTCATCCAGCGCTCGAGGCGCTGCTGCACTGCACTAAGCGAGCCAG
GATCGCTTCGAGTACGGGCTTTTCCGCGCACTGTCCAGCGCTGATCGGAGCTGAACGAGTTGGTACTGCTGGTAGCTCA
GCAGGCGGGCCATTGCGGTCAAAACAGAGGCCAAGAATCATTGATGACAGGGATGGGCGTTTCATGAGGGTCTCGCGCA
AGCTGGGGAACAACCTCTCAGACGGACAACGTTGATGACAGTAGGAAATACTGGCGGAAGGGGGAGGATTTTTGAGGGG
AGAAATCTTGCGAATCTCCGGCTAATGCTCTGTGCGAAGCTGACGAAGCTCGGGGGCTCTCTGTCGCGCATCAGGGC
GATGGGCGGCACTGTTGAGGTGGAGTAGTCAGTGAGGTGATTGCGTTTACCGGAGTGCAAAATAGGATATCCCGTCCG
TGACCAGAGGATCAAGCGGCGCGGCATGTCTCAGATGTCTGTGCGGGTGGTGGCAGAGAGTTTCCCGCTGATATTC
GGAGGTGGGGAACAATAGAACCAAAACCAACGTAAGCAGAAATGAGTGTGACCCCAACAACGCCGCCACAATGCCACT
ATCGATTGACGATACGGGTATACGTAAGCATGGAGAATCAGTATGTTACTACCCAGAGCAAGGTTAGGTTGACCA
TTGCGCTGCTTCTGAAGTTATGGGCGATAGATTATGGCCGAGCTATCAGAAAGGATGCAATGATGACAAAATCTACT
TTGATCTTCTGAACCTCGCTGCGGAGGCTCATTGCTGATACAAAAGTCTTTATCTGTGCGGCAATCTCCACAACCTGTC
CCAATACCTGGAGTTTCTTCCGAAACCGTATACGCTATGCTACGTAATAATGATTAAGTATCGGTCAACGCGATAGA
TCTGCACTTACCCAGAGCTTAAACGAGTGCCATATCAACCTGGCAGGTGCCCGGATTGCAAGTAAAGCAGATTACCAA
AAATTCATGAAACCACTGAGAACACCTGTTGAGATGGGCGTTGCTCCTAATAGCATCACCTTTGAGGAGTATCAGGCC
ACCAATAATCAGCAAGCCATCAACATGGTTCAAGATATGACAGGATGGAGACAAAGGTGAGAAGGTGGAGGCCCTCCAGGC
CAATATGCAAGTTCTGTATGGACAGGAGATAAATACTGATTTATCGCTCGTAATGAACCTGCTGCTGGGCGAGAGCGA
AAACCGTCGCAATAGTTACGGGCGATATCACCATCGGGTACGGCTTCGATACCTTCTGTCATGAAGCGTCCGAGCTAAC
TCTTTGAATCTTGTGGTTCTACGCGACAGAAGGTATTACCTGCATGCACTATCAACGTCGACCCAGGCTTCTGGAG
CGTCTATGCTTGTGGGACAAAGTCTCAGCGATGACGATGGGCTATTACTCTTTAGTGCCAAAGCGCGAGCTGTTGTTT
AACGATAGCAAGCAACAGTTTTCAGGTAAGTGGGCTACCCAGCTATCAAAACGGTTGCGCTTGATCTATAT
TATCAATATGGGCGAGCTGGTAATTTTCCAAAATTTCAACAAGCTATAAATAGCCATGATTGGCCGCGAGTCATCCATGA
ACTTAGAACTGGAATGGTGTACCGAATGATCCTCTCAGTTCAATACAAACGATTGGAAGAGCGAGCCAGTATCTGG
CAATATCTTCAACTATGAGCAATGAGGATATGGATATGAACAACAGTGAAGCAACGAGATCAATATTTACCA
AAATCCGGGCGAGTCTATTTCCGGTCTCTACAAGGGCGTGGCTAACAGTGCTCTCTGGCCAGCCATTTCCAGAGGTAC
AGCTTGTGGAGGCTTGGGATATCCCTCTCGTACTCCATCCGGAGTTTGTGCTAACGGAGATGTCTCGAAATCGATAAG
GAGTACGGAACGATCCTTGCTGCTGAGTCAGCTCAGGTATCTGCTTCAACTCCAAATGGCTCAAGACAAGGCTAAGGC
GTGCGGGGAGGTTACAGCCTTGATCAGTTCTGTCTCTCCTCAATCTCAATACCATTAAGAGTCGTATGGTGCTAATTATC
TAAACCTGCTGAACAATCACCGAACCGATACCCGACTAGCGTCGGAGTTGAGATCATGTGAGTGGCAGTCCGAACCA
GATTCTGGAATCGAGGTCTCTTACGGTGCCAGTCTCGGCGCTCAACTCAATCAACAATTCAGGCGATGAATCTGCTG
CAGTCTCAACAGTGTGCTCACTCAGGGAATCGGTGTGAAGCTTTCTCAGCTGAATATTGGCCTGCTTACACAACATAG
CCACTGGTATTCTGTTATACAACCGAGTGGCGATAACGTTGGGCTATTGGGCCAGGTTTAGAGATGGTTGAACATTAGG
GAGAAAAGGATAGGTGGCCAGATCAATCTGCTGCTTCTGCTGCTTCTGCTCATATTGCGGCCCAACAATGACCAAGTGC
AAATACCAGCAATAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGGCCGAAAGGATGAGGCTCAGGCA
TTTCTTATCCATATGATGGGCTAATGGGGTACATGCGAGCTTAAAGGCACTCTTAGAGATAATAAAGGAGTGGTGCA
TTCCGCGTATCTCTGCTTATGTTGCTAGTATCACCCAGACAGGGCTGCTGTGAGTAAAGCGTCAACGTTGAGGAAA
ATTCCGCAAGCTGCTCTGATGGTGGGCTTCTAGAGTTTGGTTGAGGAGCCTATAGCGTCCCGCCCAAGGGCTGCTAC
TATATGCAAAATAGGAAAAAACTCGACAAATGTTGGGTATGATTGAGAAAAAGAGCTTGATGCCGCTCGCGCCTTATC
AAATAAGCTTTGTGAGACTGCGCAACCGAGCTAGCCTATCTGCCAAGATATCTTGACGAACACACTTGCCATGATCA
GTGCTGAAAAGGAGAGCAATGCTCGCTGTTGGAGTATGCCCATCGGGTGCAAAAGCAATTCCTGTAAGAGATGACGGC
CAACCGGCTGAAGACTGCTCCCGCGGAGCAGCTTTGCTATGGAACAACGCGCCAAAGGCTGATGCTGTCTGAGCG
ATGACGACGAGGAAATAATGACGCGTGTAGTAGGCGCTAGTTCATAGGAGCGATCTTTCCGACTCTCCCGCGCA
TGAGTCACTCGTTCGTGCGGAAGTGTGGTAGAGCGTTTGGCGACTGATGTTGAATTCGCGGCGAGCTGAGCCTTGGGC
TCGCCGGCGCTGCCCGCAGCCCCAGGGTAGCACCTGCTTGTGCGACAAGGCCCTTTTGGCGCCCGGTACGCGCAGAGC
TTGATATCTCGGCTACCGCGCTGGTGTGCTGCTGCTGCTCAGCAAGCCAGCGCGGCTTTTTCAGGGTAAGCGT
ACGGCCAATACACCTTTACTTAGGTTGAGCCGTATCAGCAAAAGAAATCGATTATATGTGTGAGGTACCGCTCATGCA
ATTGGCTTTGGTACCGACATCTACTCGGAGACCTCGTGCTTGGTGAACACCATGTGCCAACCATCTTCTGATGTTT
CCAAACAGAGCTACGCATTGATAACGACTCATCGCCATTAGCATTTTGTATGACGCGCATCGGTACGTTGCTAACGCGAC
CATCTGCAAGCACACGAGTTTGAATCCTCGATTGCTCTTTGATCCAAGTCTGGGATTTTGTAGTCCACGATCACTCC
GCTTTGCTCCAGCTTTTGCCGATAGCTCCAAATTCACGAAGTCTTGGTAAGCAGCTGTTTAAAGAGTGCATATTCTT
TCTCGATTCCGGCTTAGCAACAGCTGTTCAAGCGCTAAATCTCCAGGCAATATCCATTTCCCTCACCTACCAAGTAT
GATTATATATGTTTACAGGGGCAATAGATCAGATAAGCTAGATCTGTACCACTGATGCGGCGAGCAATCGGATGATG
TCTTTGACCGTAACGTGCGCAGGCGAGTTAGCACGCTCTTCAAGTAGGCATACGATCATGCCCGTTTATGCGAGCGGA
CTGGATCAGGCTCATGATAGCTGCTGCTGTTTGGCACTGCGCAGCGATCCGGCAAAATTAACAGTTGAGCGCTCCGAGCG
CCACGGCGGTATCTGGTTCTCCACCTGATTGTTGTGATGGGACAGCCCCATCGTCCAGGTAGCGGCTCAGCGCTCCC
CAGCGTTTTCAGGCTGTAGTCCAGGGCTTGTGCTGCGGACGCGCTTGGGCAACAGGTGCGCTGGGCGCAACTGCGT
ATGCAAGTGTGCTGATCGGTACCGCATTTCTGACGATTCGCAACGGTCTTCTGTTGCTCATGTCCGAGCGCTGGC
GTTCAACCTCGTACAAACCGCAATGAGTGACGCGCTGTTTGGCCAGGTGGCTTTTGTAGCGACATGAGGTGGAAG
AACTTGGCGGAGCATGAGCCATGACGCGATTTCAGTGATGCTTGTTCAAAACCGGCTTGTATCCAGCGAAGTCTGTC
GCAGACGAGCTTGCATTCCAGTGCCTAGGAAGTTGCGTGATGTTTCTCGGCGAGGCTGGGCTGAAGTGCATAACCA
CCGCTTTGAGCGCCGAAACGGCGTGTGCTGTAGGCCAGACATAGACCCGGTGGGTTTCTTCTCGCTGGTGCAAGC
ATTTGACCGGTTGTTTATCGGCGTGGATCAGTCTGGTTTACGACGCGCTTACGCACTGATCGACAGTGGCTGAAG

FIGURE 2G

CCGCACGCCAGTTTGTCCGACCCACTGCGCCAGGGTCGAGCGGCAATTGGCAGCCCGCGCGCCAAAGATTTTCTCT
GTCTGTACAGCGGCAAGTGATCGGCAAACTTGGCCACCATCAGTGGGCCAAACAACTGCGGTGCGGATGCCTTTATCA
ATAACCTGGGCTGGCACCGGCGCTGGATCAGGGTTTCGCACTGACGGCAGGCCATTTGCCCTCACATGTTGCTCGAC
GGTAAACACGCCCGCGTGTAACTCAGCTTCTCGCTGACGCTTCGCGGATGCGTTGAAGTTGGCAGCCGAGGCGCATT
GGGTATTTTCAGGTTCTGGCGCAATCACCTGCGCGGGAACCTGCGCGGCAATGGCGAACGTTTCGGGGATTGCGGTGC
TCGGCTTGTGGCAAGCTGGAAGGAGTTGTTTCAGCTCGGCTCGATAGCTTCAAGGTGCGTGTGAGCAGAGCATCCAG
CAAGCTGCTTGCGCCGAATGATTGCTCGCTGCGCTTGGCAAACTTGTGGCGTTTGGAGCAGAGCATTCGAACCTGA
ACTGTTTCGATGAGGGTTTCAATTGTTGCGGATTTCTGCTCATCGCTCGACCTGGGATTGCAACTGCAACGCTGTGCC
GCCAAGGCACGAAGCTGTTCCGGGGTTCATCTGCTCAAGGTTGGGCGAGGAAGTCATGCCCGGATTGTGCCAGAGCGC
GCGATGCGACGCAAGTAGACAGCCAAATGGCCGCGTTCACAGCATGCTGATCACACCGCTGTGCCGACGCGCTGCC
ACGGCAGGCCACGACCAAGGCTGGAGTTGTTGCTGTCGAGTTTCGACCTCGCAGCCATGGCGAATGCCGGGCCAGTGG
AACTTGCCTTGGTTTCAGTCGACGCGCGGCAAGCCAGATGCCACGCCATCGTGACCCAGCACTTTCATTGCGTTAGCCCG
GCGATTGGCGAAGACATAAGCGCAGTGCAGCTTCCGCGCACCGAACACCGCAATTACCCGGGCTAATGCCGTTCGCTGC
CGGCGCGCATGTCCATCGGTTTCGGTGGCTAGCCAGATCGCATCGATGCGGATCATCGAGCGAAAGCCCGGATGAATTGCG
CGCAGCCCTCGGGATCTGAGGCTGGCTATTTACCGTGATCATTGCGCCGGCCATGGGCGATTCAATGAGCACTGACGTT
TCGGCTGCGCGTTTGGGGTGGCTTCAGCGGACAAAGCTGTAACGAGGCTGGTACGAGGCTGGTACGGGCTGGTCTCGATAAAGCGT
CAGCCATTTGCGGATGACATTGGCGTTGATGCCGTGCTGATGGCGACACTGGACACCGTTGCCCGAGTTGCGGCAATT
CCTGAACGACCTGGGCTTTGAACGCTTCGGGTAAAGAGCTTCGTTGGCGCATGGAATCCTGGCGATAAGGGTGATCGCG
TCCGCTTAAAAATACCGGACACCATCGCCCTTAATGCTGGAGTTTCGGAAGGTGAGTTTCGCCGAGCGCTTACTTTTTCAGG
AGACTACGAGTCATGTGCGGATGCACTAATGACTTGCCAAATGAAGACAGTCAAGATTGACGCTTTAGTAATAACGGAAA
TTATCATGACATAGGTTCTAAAAGAGGCTCCATTGGAGCCTCGTATTTTGTATGGTCGGCGTAGCACTTACTTTAGTGA
TACATGCTTACCCATGAAGTGTATTCGTTAGATAAGGAGCTGACTTTAAGTGAGTTCTGTAGCCATAGAATCCCCAA
GACCATCCCCGTCGGGAATTCAGTGATTCCATTGCAATGACCGTATGGCGCCTAAACCTAGAAGCTTGGGAGCATAG
GGAATGCTCCATCTGAACACATAAAGGGTGGGTGAGCGATGGTCTCTATCTTCTTCCAAGAAAAGTACGAGGTATAAC
TCCTGCTAGTACCTCGTAGCCGATCATTCTGGAGAAGAGTACGACATACGGCCGCCATCGACTCTCATGCTCACATG
AAATGCTCTGTGGAAGTGTGATTTCTTAATTTTTTATGAGATAAAGTAAGGTTTCTCTGAGCTTTTTCGAGAGGA
GGTCGGCGTATATCAACTTTAGATATTTCCGACAGGAGGATGGCTGCATCCATAAGGGAGGGTATGCGCCTGAAATGAG
TTGATAGATCGTGACTTTATGGAGAGGAAGGGACGCAAGGGAAGTTGGAGGAGCCAAATACCTGATTACGAGACTTTG
CATGCTCATATGCACTGATCTCTAAGTGAGTTGACCATCTTTATACGTATTTCTCTTGGATGTCGTTTGGCAATAAA
CTGCAATTGACTCGATCCAGCAACGAGCCTCGGTACTAAGCGAAACGGGAAGGTTTCGGTCTCGCAAGTCAAGAAATCG
AAATTTATTTAGAGTTTCTTCGACTTCTAATTTTCTATTGCGTGAATATATGGGTTGTTGGATTGAGGCTCTCTTTTG
TTTCGATGAGGATTAGTCCGGGTAGTTTCTTAAAGCTCATTGCTCAAGAAAAGGACTAAGGCTGGAAGTATAGAGTGC
TTTGGGTTTCCAGTAAAAACGACGGGGCATTCTCGGTATATTCAAGGTCTAAACAAATATAGCGCTTCCAGTACCCCTG
GTGTTATCGTTGCGGCCACGGATCACCCAGAAATGGAGCGCTTCTATTATGCAATGTGCTGCTCCAGTTTATCGACAT
AGCGAGTGCAGTGCAGGAGCCGAGCAAACTGCTACCGCATTTGGCAAACTTGAAGTGGATATTCTGTTCCAAACACAAGA
AAGGAGTTCTTCTGCAAGCGTTTGGCATCGGTTTCAACGCCGAGATGTTTCGTGAGGCGGACGAGTACGAGTACGAAT
AAGAGATAAAGCCATGTTTTCCCTTGTGCGCTTGTGGCACTTACGAGGTAAGGACTGGCTCGCTGTGGCACTTAGTG
CCAAATTTGCGAGGTTTATATAGTGACAGTGGTCCCTCTCGGGTTTCATGCGCGCACAGCGCGGCCAAGACTGTACC
CCATTTATAGATGACGCAAAATCTTGATTTTATCTGCACTTCAGTGCTATGTGGTGAGCGTAGAGTTTATTTTC
AGCTGTCTCTGTGCTGTTGCTAGGAGATGCAAGTAGTTTACAGCAAGATGTGGAGAAGGACTCCAGAAAAAGGCCACC
CGAAGGTGGCTATGAGTTTCATGCTGCGATGAGTTGTGCTGCCAACGCCATCTCCACAGAAATCCCATCTGTTCAACG
AGTCGAGACTGCAGGAGCCGAGCAAGCTCTCCGACGTGCTCTGAGCCACAGGATCTTCTGGCCATCAGCTGTAAGCAGGTG
ATCTGCGAGCTCGCGCGTAGCCGAAGAACACACCTCACCGGGTCTTCTGCCGATCCGCGCACGACCGCGCGCGC
CTGCTGACGGGTGTACAGTTGTAGCCGTTTTCAGGAACGCGATGGTGGGAAGTTCGAGCAAGTCCAGCCCGGTCTTCA
CCAGCTCCGGGTTGGTGATCAGCATCATGATGCCGCGATCGACCTGGTTCGAGGATCCAACTCTCGCGTCGAGCGGTATCG
ACCGAAGCAGCTAGCACTGCACTGCACTTCCAGCCGGATTGCTCGAGCACTTCTTCAGCCTGGACGTGGTGTGCGCGCTCC
GCTGTAGACGGTGTATGCCAGAACCTTTCGGCCCTTTCGCTTCTCTCGAAGCAGAGGTCCACAGCACTGCTCCTTGG
GTATCAACTGCTCGTACCGAAGATCGTGGCAGCAAGGCCAGTGTGTCGCGGTTTCGCGGATGCTTGACGATCTCCGCT
CGGAACAGCAGAGTCCGGCAAGCCAGCAGCAGCTTGAGGACACGCCAGGAGCGTGGTATCTCGTCGCGCCAGAGCCTG
GCGGAGCTCCGCTGTGAGCGTGGCGCGCAGGCGCTGATAGGCGCAAGCCTGCTCAGGCGCCATGGGCACTGCGATGAAT
CCTCTGCTAGTTCGGGCACTAGTTCGACCAATATCTTGAGCTTCAGGAACCGGTGAACGGCAATACGAAGCGGTGG
ATGCCCTTTGGGCGAAGCCGGGAGCCTTACCGTGCCTACCGAGAGCTTCTGCCCCGCGCTGTCTGTGCGAATCACC
GTGCGCTCGGTGTAGATATCTTGAGCACACCGTGGTGGCGCATGAACGACATGGCTGCGGAGCCATGCTGCGCGCGC
CGTTGGGCGCATAGCCGCTCTCGATCATGCGCTGGGTGAGGATGCGGAACAGGAGATAGAACAGATCGTCGGCGTAGCCG
CCCATGAGCGTTCCGGTCAGCAGCAGGTTTTCGCTGCTTGGCTGCGGAACGCCCATGGCCTGGCCCTGGGCGAGCC
GCTGTTCTGTACTCATGCTCCTCGTCCAGCAGCAGGTCGAAGTAGCCATCAGGTAGGTAGCGCTTGATGAATCTCGG
TCCGTTGGTAGCCGCTTACCGAAGCCGAACCTCGATGTTTGGCATCGATCGCTCCATGCGTTTGGCCTGCCGATCGCTG
AAGACGAAGTTGCCCTTGGCGTCCATCAGGTTGATGAATCCGAGACGTTGTCCACCAACATCGTGGCCAGGAAGTCTCTC
GCCGAAGTCTGTACAGGCGCTCCGCGCTGACCGGGCCGATGGTTGGTATCCGGCACATCGACTTGAGAACTCGTTGCGC
GCCGTTTGGCGCGTCCGGCTTGCCTGGTGGATCAGCGTCCAGAGCGCCCCACGGCAGGAGGAACAGGTACGTCGACGG
TCACACGCTCGAATCTCTCCACGCTGACAGGTTGCCCTTCCAGGTCTCGAGGACCTGTCCGCAATCCGGGACGCGC
GAGCAGTTGGCGCGCGCGCGCGCTTCTTCCAGCAGGCGAGCGCCAGTGGAAACCCATCCGCATCCGCAAGCGCGCGA
GGATGAAGAACTCTTGGCGCCGCTGTAGGCGTCCGCCATCTGATCTCGCAGCTTGAGCAGCTTGAGTAGAGTATCTGGG
CCATTGAGTACCCAGACGCGGGCGGCTGGGATGGTCTCAGGATCTCGCGCGCCACTTGTAGACAGGTTGCGGCGGAGA
GACGACAGGTTCCGGCGATAGCCGGCGGCTGTCAGGCTGCGACAGCGATGGCCATCATGGTTTTCCCGGTGCCCA
TCTCGGCGTTGATGATTCCGGCTGCTGTTACGGTCCAGCAGCAGGCGGTGATGGCCTGGACGACCTCGGCTGGGCC
GCCAAGGGCTTGGCTTGGTCCATCACCAGTGGCGGTGAGCGTTGACGAGCGCGTATAGACGGGGGGTGGGA
GCGATTGAGCACTCCAGGAGCTCGTCGCGCAACTCGTCATGAAGTCGGTCAGGTTGATGTTTCCAGTGGGAGCGCGGA
GGGCGCGCGGTTGGGTGAGAGGTTTATGTGATCTCTGCAAGTGAAGAACGAGCAGGCGTGCACGCTCCCTTTGG
GTATCGGCATACCTGCAAGGTTGGAGTGGAAAGGAATCGCCGCGGGGGGAGGTGAGTAATCGCTAGGAGCAAAAGCGTG
GTTGAGCTGCGGTCTGCCTCAGTGATGATCCAGAGCCTGGATTGCTGCGCGCGTTCATGTCGTAGGACGACAGCAGCCG

FIGURE 2H

ATCGCCGAACCTTCAGGGCTCTCTGGTTGGTGTTCAGTCTTCTGGTCCAGCGTTCCCAATCTCCCGCCAGATGGCGTT
TCAGCAAAGGCGCTGGGTTGAGTTGACCGGTTTCACCAAGTGGCTACGGCTGCGGTCATTACGACCCGCGCCAGGGCTG
ATCAGCAGCGGTGCGGGTGCAGGGGCGGGCTCAGGTTGGGGCTCTTACGCTTCTTCCGTGGTGCAGGCGGAGGAGCTGAT
GGTCAGCAGCGGGCTGATTGACCGAGGAGGTGTCTTTCCATGCCCCGATGATCGGGATGAAACGGTTCGGTGAGGA
TCCTCACCTCGGTGATGTTGCCGTGCTGCTCTCGGTGAATTCCGTCTTGCAGACCTTGTCTTGTAGGTGTACCCCTTC
ACGACCAGGATCCGCGCCGACTTCGATCGCACGACCCAGATATCGCGCCGCGCGCCAGGGCCAGGGCCAGGTGCCAGCG
AGACAGCTCGCGGACTGGTGGGCGCGGCTGCAGCCCGCTTGCAGCAAGTGCAGGTGAAAGTCAGGCCAGAGACCTCGCA
GCCGCTGGAATTTACCCGCGCAACTGCTCCGGCTCCAGGGTACTCGGTAGAAGTGTCCAGCTCGCTGGTGGCGCGCAGA
ACCAGTAGGGTTCCACCGCCAAGCCGCTGGAATTTCTCGGGCTTTTCTGGCCGCTCCGATCGCTCGAGGCGAGA
CCTCACCTGATTGGCGTCCGCCGCGGCGAGTCTGCGCGACGACCCGGATGCCGAAGATCACCACCTGCTTGAAGGTAG
GATCCGCGGCTGCTAGATGCGCAGGCGGTGAAGTGGTGTCTCAACAGCCAGTCAGCTCATCGTCCAAGACGTAGTGA
GGAACAATCAGAACCATGACGCGCGCTACTGCAGCAACGGCAGGCGAGCTGGTAGAACGCTTTCTCCAGACGCGCGG
GCCGCTGCCCTGGTACTGCGACGACCGAGTGGTCCGCCACCAAGGTGCCATAAGGCGGGTTGAGCCAGAGCAGTCCGA
ACGACTGCCCTGCTGATCATGGTGTGAAAAGGTCACTGTGCAGCACTCGGTCAAGCAATCTCGGGCATGGTCCGCGCGC
TCGCGGTGCTACTGCAGACGAGGGCTTGGACCTGATCGCGGCGAGGGTGTGTGCTGCCCTCAGCCAGGGCAACACCCCT
ACCGCGCAGGGGTACAGATCCTCATCTTCCGACGCGGCAAGAGTGAAGGCTGCAGAGCGCGTTCAGGGGTGACCT
CATCGGTAGGGAAGTAGCGTGTGCTGCAAGTTCGCGCGCAAGCGCGGGAACATGAGGGCATGGGGCTCTTGTAGTTC
TGAGAGGATGGCTCAGGCTTGGCGCTGAGCATCGGTGGTGAAGATGCTTTCGCGGATTACCTCACCAGGGGTGGGCTCA
AGCAGTCGACCCGAGGGCGAGTGCAGGCGCAGACCTTGGCATGCTCCGGGTAGGGCCGTCAACATCTGGTGTGCT
GGTGAGAACCTCCATCAGCGCTCGCGCCAGTGTGCGAGGAGGGGAGCGGACAGGTTTCCATCACCAGCGGCCAGAGCC
TCTGGTGTGATCTCATCGCGCTGCAGGAGTGCAGAGGCGAGGTGATTGCTCGGTGCGGCGCCGAGGCGCGCCGATCG
AACAGCCAAAGATTGAGCAGGCTGCCGAACAACGTGCCGCGGAAGTGCAGGGTGGTGGCTTCTCCAGGAGATCTGGTT
TGGGAAGACAGGTAAAGCGGCGCGCTCAGCATGATGTGGAATGGTGCATGCCATTTCTCCGCGCCAGCGTCAAGC
TGGCCAGGAACCTTGTGTCAAGGTGTGCGGCGCCAGGCGGAAAGAACAGGAGTGCAGTGTCTGTCGACACGCGAG
GCGTCGACGTACAGGTCTGGACACTCTTCGATCTGGTAGAGCGGGTGGGACTGGGCATAGGAACCTCTGGAAGGAGGA
GCCAGCGCGCTCAAGGGGCGGTGAAGCCCTCGGGGTGTAGTCCAAGTGGTGCAGGGGAGCGTGGTGGTACCGAAGC
CTCTGGTGTAGGGTCAAGCTGAGTCCATCGCTTGGTGCAGCGCCCTTGGCGCATCAGGAAGACCTGGCAGAGGTA
CTGATCGGTAGCGCAGCAGCGCTTCCGCTGCTCAAGGTGAGCAGAGCTCTCGATGAACAAGTGCATGAGCTCGT
CTGCGCGCGCAGCTGTTGTTGGACAGTGGTCTCGAGGTAGGCGAGGGTTCTCGGTGAGGTGGTGAACAATGGA
TTATGGTGTGCTCTGGTGGTAGACGAGGAGCGCCCTGAGGAGCCCTCAAGGAAGTGGTGGTGGTGGTGGTGGTGGT
GGTGTCTGCTCAGGCGTTTACGCCATCTCTGGTGTGATCTCATCGTGAGCCAGTCCAGGAACCTCCAGCGGACAGCG
TAGTAGAACGGCGCGCGCTTCCGGAATGGTCTTGTAGCCCCATTTGCCCGCGGAGCAGTCCAGCAAGTGCAGCTCGAT
CATGGTGTGATGAGTACCGGCGCGCTTGGCGAAATAGAATCCGGCAAGGTGAAAGTTCGTTTTACGACAGTCCAAAGT
CATTGCCGAGACTGCGGTGCGCAGACCTCGGTGCTGCTCGCAAGGTACTGGTGGGGCCAGCAGTCCAGCAGCAGG
TCTTCTCTGCTGATGTGAGAAAAGCCATCCCATAGGTGTGCTCTGTGGTGGCCAGGTCTCCTGACGGGAGAGCC
GGCAGGGGGTGGTAGGGTGTGCTTCAAGGTGTGCGCGCTCAAGGCGCTGGAGGAGCCGTAAGCTGTTGGCTGGCT
GCCAGTCGAGTCAACAGGGCTCGTGTGCGCGCATGCTCCGCGCAAGGGGTCTTACCCGCTCAGCAGTGCAGT
ATGGCTGGTGCAGCAGCGATCGTGCAGCACTGCCGCGCGGAAGTGGTGCAGCAGATTGTCCAGTCCGACCTCGCGCAG
CGAGGCGCAACAGCAACAACCTCTCGCTGAAGCCTGGGGTGTGTAGATCGATTTATGCGCGCTCTTGGGGTCCA
CACCTGTTGTTGAAGTCAGCTCATAGCCAGCGCGCGAGTGCAGCAGCTGTTGGCGTAGTCTGCTGCGGTGACCG
TGGTGTCCAGCTTGACGATTTGCGCTAGCGCCAGACCGTCCCGAAGCTTCTGCATCCGCGTCTGTGTCAGAGAGCG
GCCTGCTCTGAGGAGCGGAGTGTCCATACCGAAAGGCGCGTCCGTGGTGCAGCAGGACAGAGGCTTCGACTTCTG
GGGTGACGGGCACTTGGTGGTATCGCCGCTACTACTGTGGGAGTGGATCTCGGGCACTTCGTGAGCGGATCCACT
CATTGCCGGAAGACGCTCTACCTCTCTGCTGCTGAGGCTGTCCATATCGTTACGCGTATGCTGTCCACTATGGCGCG
ATCTCGATGACTAGCCAGCGCGCGGAGTGGTGGGGCGGATATTGGCGATGACGAAATCACCAGCGTACTTGGC
TTCGTTGTGTTGATCCAGGAAGGCATCTTGTATGATGAATCCCGCATCGAAGTCCAGAGTTTCCACATTGAAGTCCG
CATAGCGACGTTGATGGTTCGAATGGCCAACGTGTCGGAAGTGTGATCATGGTATGACCTGCGAAAAGGATGATCCCG
CCGAAGCGGATCGTGGTGAAGGATGAAGTACTCGGTAAACGGGTGTTGGCGAGGCACTCTTCTGCAACTCGATGCG
CAGGTAGGAAGGCTCATCGGACGGCTCGACTATCGAGAGAACAGCAGCGCGAGCGCGGCTCTGCCCTCTTCTCTCT
TGTGGTACAGGTGGAATTGGATGTGCTTCCGCGTCCGCTGGAGAGAAAGCTCTGCCAGGCGGTTGAACGACACTGGCC
AGCCGGGTGCCATCTCTGTCAGCCCTGGCAGGTTGGCGAGATAGACCAGGCGCAGGCGCGCTGGTGAGAACGAT
CGGAGTTTGAAGTACGACAGTGTGACCTGCGCTGGAGTGGTGTGAGAGAGGTGATGAGAGGTCTCCGAATACGCA
GGCTCTCCCGCGGGATGAAGTCCCGCGTGGGTGGTAGGATGCACTCCATCACTTGGCTGATGCGACCGCGCTGCCG
AGTGTGTGGCGGCTTGGGTGGGTAGCGCTTGGCGCGCGGAGCTTAAACCTGGCTCTTGGGAGGCGCTGTCAACG
ACAGCGAGCAATTCGAATTCCGGGGCTGGCGGGATTTCCCGCGCGCGGATTGCTCAATCGTCATAAAGCGCGAGGCC
TTGAGCGTGGGCGCAATGGATCAAGTACAGGAGCGGAATGTCTGCTTGTCCCGCAGGTGAAGCAGATCGACAAGGTG
TTCGTCAGGCGCAAGGCAAGCTGTTCTGTGCAAGAGCGGCTGCGGTCTCTCAGGATCTCCAGACACTCATCCGTC
AAGGGGTGTTGTGAAGCAAGCCACGCGACCGTGTCTGTTACCAAGAGAGCGATACCAACACCGAGGGCAAGCGG
ATTGGCATCTTGGTCGAGAACCTTTCTGAGAGGCTATTCCAGCGCTGATGGGTAGATGTGTTTTCTGATCTCCAGCA
GGGTTTATAACCGCGCGCAACGAAAGGTATCGAACGCACTCTCGGCTCGGTGAGTATGATCTCCGATAAGCT
TTTTCTGCGCTCTGCTCAGCTAGATGCTGTAAGTACGTTTCTCACTATCTTGGTGAAGGCAATTTTCAACTCTT
TCATGCTCAACTCTTGGACTACACGCGGAAATCGTTATTGAAGTAGCAGCTCGCACTGAAAATATTCTGCTGGAG
GTGCTGCTGGGTTTATGACAGAGCGCGTAGCAGGTAGATTGTGCTTCTGTCAGTGTGTCAGGTTTCTGCTGATGC
GGAACCTCTATAGCCGCAATAAAGGGTGTGAGTGTGATGATGATCTCCGTTGCCGGGAAGCCCATGCCCTTGG
AGGCTATGGACCTCCAGGTGAGTAAAGGGGCGCGAAGCGCCAGGTGATGTGAGTGTGTCAGTTGTTGTCAGTTAAACG
TCTTCTTACGCGAGCCTCTGGTATCGCTCTACAGGCTAGTATTGAAGTACCAGGTGATGTGCTGAGTGCAGTAC
ATGCTCAGCTCGGCTGCGCAAGCTTCGATTGCATCTGAATTACATAGAGCTCGATACTGAGTCCGCGCATCGCCAA
ATTCATCAACCGGCTTGGCACTGGATGTAAGTCAAGCTGCTTGTGACAGTGCAGTCAATGTTGAGTAGAGCTGGG
AGCCTTCTGGATATGTGTTTATGGTGAAGTCAATTCGAATCAAACTCGGAGAGGTACCTCCAGCAGGGGTGATGAC
CCGCTCTGGGTGAATAGAAGTGCACATGGGCAAAAGGCTTAATATTCCAGTTGCGTCCAAGTGGCTGGTCT
TAATGCTGTGGGCAAGTTCGCGTGAGGGGTAATATTCAACGACTCGCGGAGGCTGGTCTCTCTGTCAGAAAGTGC

Figure 2

FIGURE 2I

CGATGTAGTAAGCACGGTTCCTTATTTCATAAGTACCTGCAACGGCAGTTCAGGCTGAGATGGGCGAGCGCCAGCTTTCCG
TATTGCATGGTGATTCTCCTTTTCGGTAGGAAACACCACTGCCATGGGGACTGAGGTCCCTAAATGGGAAATGTTTGGCG
ATTATGCCAGGGCAGATACCTGGGCATTGTTGTTGGCCAGGAACAGCCAGGCGAAGTTGCAGAACTGACTGCTGAGCAGC
AGCTCGTCTCGACGTGCTGAGGCTAACCTCCTCAACTTCGGGCAGGGCGGGCAGTTCTACGATTGGATGTTGCTCAT
GATGATCTCCAGATAATGGAGAACACCACGCCCTGGGGGAGATGGTGCCCCCGGATGGGTATGCGATGAAGAGTGGCG
CATAGCGTAGGTAGTGAAGGCCCTCGGGATGAGGGCCGGTATTCAATCAAGCCGTGTTCAAGCAGCGGCTGTTTCAGCG
TGCTCAGCGGGGGTGGCGGCTCACCTTGAGGTTCCGCTTGGCTTGGCTGGGCTTTTCAGCTTCTTCTTCGCATCGTA
GACGGTGGTGCCATCCACTTTCATCCAGGAGATGAAGAGCAGCCGGCTTTGAGGCTTGGCTGGGCTTGGCTTGTCTTCT
CGCCTTCTGGTGGATGAAGGATCCGCCAGATATCGCCGATACGGAAGGAAATCAGAACCTTCTTCTGGCCTCGACT
GCTTCTTTCGCAACGGCGGACAGCTTTTCAGCCTGGGCGCCGACCACTTTCAGTTCGATGATAGGAGTATTCACGCTGTC
TGCCGCGCATGGAGGGTGGACGGTTACGGCGAGGAATGTTTACCTCGCGGATCGGTACCTCGCGGATGCGATTGA
GGTAGCCGATACCGGTGGTGTGAGGTGCAAACTTGGCTTCTTGGGCTTGGGTGTTGTTGCTCATGATGATCTCCAG
TTTCGTAAGCGAAAGCGGAGAAACACCTTGCCCGTCAGGGAAGATATCCCGCATGGGTGGTATGGGAGGTAAATC
AGAGGGAGAGGGCTTCATCAACCTGTGGGTGATTGATCGCGCTACCGGAAGCTGGTGGCATCTTGGGTTTCGATCTACA
CGGAAGTTCGTTGCTAGTCTTTCAGATCATGACTACTGGGGAGGGTTGCCGACAACGTGACGAGCTCAGTGGTAAAG
GATGGGCAGGGGAAAAATGAATGTCAATACCCATGAATGAAAAAGCGCCAGGGTGGGCGCTTTCATGGTCAATGAT
GACGGGGAAGAATCCGATGTTTTCGACACAGGTCTTGACCTGAGGCCAATGAACAGAGGGCTATGCCGAGCAGATA
ACCACAGGCTGGTTCATGATGATCATGCCGAACCTCGGTCTGGGACATTTCACTCCTCCTCGGTGAGAGCGATACCTATGA
GCCGATTTACCATACCATGACGAAAGCATATGCCGTGACGAGCGGTGCTCGTGAGGAAGATGCCACAGGCTCGCC
GCAACGGCAGTGGTACCGATTTTCGGAGATCGTCTGCGAATGCAGAACGTTTCGATTTTCATGGTTTTCTCGAAAAATG
CCGGAACACCACTTCGCGGATGGTTCCCGGCGAGGGTGGGGGTGGTAGCAGTGGCTATAGCGGCCGAGAGCAGCTACT
TGTTTCATCGCTCATGGTGGACAGTAGTCTGCCAGCGTGTGGTATTGCTCCTGAGTACGCCGATGGAGTTCCTCAAGA
GACTCGCTTACAGGAGGTGCTGGCAAAAGAGCAATCTCATGATCGTTCAGGGAATCTAACTGGAGAGGGGCGCTCCCA
TTTCGCTTTCGAGTGGGTGCAATCAAATGGTTCAGGGAACCATCTTGAGCTCCTTCTGACTGGCGTTTCAGCTTTG
GCGGTGCGCTTGGCGGCTTGGGTTTGGCTGCGCGCTTGGGCTTTTGCACTCGCGCTATTGGCTAGATACCACTGATT
CAGGTTGACCTTCTTCAGCGGACGATGGCTGGATGTGCTGATGTTCTAGCTCGCCAGCGGTGGGCGGCGCAGGA
AGGCAGGATCGATGCGTGGACTCTCAACGCAACCGCGCTTGGTGTGCACTACGACCGGAGGGGCACTGAGCGCTCTG
CGGAATCAGCACAACACTGCCGAGGGCAAGCCGATTTCAAGCTGGTGGATTGCTTGGACATGGTGTCTCCTGCA
GGTTGAGGGTGGGAGAAAACTGGCCCTGGCGGGGAAGTGGATTCCCGCTGGGTGGGAGAGGAGCTTCTATCGAC
CGAGGCCGACCCGCGCTACCAAGAGCTATGCGGTCTTGGGTGGATCAACAGGCTGGCTGTCAGCATTTTCG
ATCCGAGCTGCATGGGGCATGTGCTGACAACGTGACGAGATCACGCGTTTCATCATCGAAGCGGGGGGCAAGTGTCA
AGGTCAAGCAGGAGGGGAGCGGATGCGAGCGGGCGAGAGCGAGGGCGGAAGAGCGCGCTGTTAGAAACCTGTCAAAG
GTTTCCCGACCGTGTCTGGAAGGGGAGTCAAGTGAGCCACCTGTGGTGAATCGTTCATCCCTGGGCATTGATCGTCTC
GATCCCATCACCCGGCCATTATCGTGGCTTGGCCAGCAACAGTAGGCAAGCATCCGCGTACAAAGGAGGCGCTGCGTAT
CACCGGCGGGCCACCGGCGAGGAGAAGGATCCCGATGGGTAGCCTGATAGTCTGGAGCAGGAGCATCAGGCTACCCATG
GAGAGGGGAAAAAGAGGGGCGTAACACAGTAGCAGCCCTTAAATCGAGGAAACACCGAACCTTTCGAGAACCTTATGTG
CCTTTAAACTCGTTCGACACGAGTTGACGCTGCATGACAAGCAAAAGAACAGGAGACACAGGAGAAATTCACACAG
ATCCGACGCCCCACCGGTGTTGTAGCGTGTACGGTTTGGCGATGAGGCGAACCTGCATCACAGGGCGCTACTGAT
GACCGTCACTAGCCAGCGGGTGAATCAAAGTTGATGGCATTATGTGTTCAGAAGGCCGGTTATCAAGCTGCGCCGGG
GTTTACTGGCCTGTGGAGGTTATTTGACCGCGACACCGCGCCCTCGGGGATGGGCGTAGGGAGCGCAATGACAGTGGT
CCGAGGCGCGGCTGACTAATCGGCGTTCGCGCTGACGCGAACCAAGGTGCGGATCCGTAACAGCATCACGGCCGCGC
CAGCCCCCGCGCGCGGAGGCGATGAAGGCGTAGCTGGTGGCGGCGGCGATGTAGCCCTGAAGCCGCGCTGGCTTGGGA
ACCAGGCGGGTAAACCAAGAGCGGAGTCCCGCGCGCAACCGGTGAACCGGGGTGGCGTGTGCGTACTCCGCCAGACTG
GACCGAAGAGCAGCAAGCTCCTGGCCAGGGTGTGACCTCGGCCAGTTCGGCGCTCGAGGTGCGGTGGTGGCTTGTGAC
ACTCAGCAGGTCCCGGTGATGAGCGCGAGCACAGGACGATCCACATCAGCGGCATACGGTAGTCTTGTGCTGGCGG
ACCTGCTAGTTCAGCAGGTACCGTACGTTGGTTCGAGTTCGGGGTCTTGTAGGGTTCGGGTAGTTGAAGCTGGGCC
GATCTCAGTGATTAGCCCAACAGTAGTGTGGTGCCTAATGTCTGGTTGAAGCCCACTGCGCAGAAATCGGCCGCGCAGG
TGATCGCGTAGCGGACGTAGTGTAGTTACCATGGCGCTATCTGACCATGCAACGTACACAGTCCCTGATCACGCCG
GCCGTGGTGTTCAGGCGCATGGAGGCGCGGTTTCGTCCTTCCGTTTTCGGTGGCAGGACAGGACAGTCCGGTGGT
GATGCTTCCAGCCCGTTCGGCGAGCAGTTGGCTCCTTCGGTGGCCACGCTTTGAGCTGTAGGTACTCGCCGACTTCCG
TCCGGCCCTCGGCGGTGAGTTTGGCCGCTTTCATCTCGCCGCGGCTGATAGCGTTCTTGTTCATCCAGGAGCGCACCCAG
GTGCTGTGCTCATGTACAGCGCGCGGCCATTTCTCGTTGTACAGCGCGTGTACCAAGGGTCTGAACAGGCTCC
GGTGTACGTCTCGCCGCGACGTGACAGTGCCGGCTTTCACCGTACAGTGGCTGTACGTTGCGCGCGTGTATGTCCG
CGTGGTGGTGTGTTGCGCTGGCGGTGATCGCCCGGCTGCGGCGATATTGTTGCTCCCATGTCCAGCGTGGTATTC
ATCCGGTTGAGTTCAGGATGACCCGGGACGGCATTGCGGTAGAGGTACTCGTTGGCGATGGCGCGCTTGGGAAGACAG
CGCCGTGCGCAGATGTCCAGCGCGGGAGCGCTACCGAAATGCTTAAGGCCAATGCGCAGGCGCGCCCTGGCGATAC
TGGTGTGGTTTTTCAGATGTAGCCCCGGTGGCTCCCATGAGCTGCGCGATCCGGCGGATCGAGAGTTCCGAAGCTACC
TGTCACCCCGTGGTACGATCAGCGTTTCGAGCTGGTGGCGCGCGCTTGGCGGCCAGGACCTGGTATTGCTGGCGGTA
GATGTTGGTGTGCGGAAGCTGCGGGCAGGTAACGGGTGTGCGCAGCATCGGACCGGTGATCACCGCGGGGCGGTGG
CGCGCGCTGCGCAGAACCGTGTGAAGTTGCTCTCAGGTACTTCTCGCTGCTTGGCCACCTGTGCTGTGCTGG
GCCGCGTGTGCTGCTCCAAGCGTCCAGGTAGCTCATCAGGACCGATATGCCACCGCGGCTGCGATGGCGATCACGAC
GAGGGCGATCATCAGTTCGATCGAGATGAATCACTGCTGCGCGTACTCCTCATGGTGTATCTCCATAGGCTTTGCCGG
GACTGGCGAGCTTGGTGTGCTGCTGACGAGATGAGCTCTTTCGGCGGATCTTGTGCGCATGGAAGGGCGGTGCTGCGC
AGTGAACAGCAGTGGCTCGGCGGTGAGGCGCGCGAGGAAACCATCGCTCTCCAGCGCTTGGCAGATCACAGCTTGGATG
CTTCGGCAACCACTGCTAGGCGTTTCTCGCCAGTGGCTGGGCGAGCGTACCAAGCGCTGATCGCTGATGACAGCTT
TTCGCGTGAATGGTGGCGATCACAGGCTGCCGCTCAGAGAGGCTTGCAGGCTGGTAGGCGGTGCTCCTCGCGGAT
TTCGCAATCAGCACAGGTGCGGCGCGCGCGCAGCGTGGCAGCAGCGGCTGCTGTATCCGCTGAGCGCGGTGAGG
TTCTGACCTGGATGACGCGCGGAGCCATGCTGACCGCTGAGGTTGGTTTCTCGGGGTCTTCGAGCGCGAGCCACC
CCGCCCCAATCTTCAGGCGGGGCGAGGACAGAGAGCGCGGAGCTGGTCTTGGCTGTGCGCATCTCGCGCAGAACAG
GACCAGGCGCTGCAACTGCGGATCCATCAGTGCGGAAACCACTTCGCTCGGATAGCCGATCTCTTGAACCTCCCGCAAT
GGGCGCTGACCGCGCAGCAAGACGCTGCCACTGAAAGCGCTTTCGAGGAGGGTACGCGAAGCAGAACCGCGTGC

Figure 2

FIGURE 2J

ACGACGATGGCGAACTCAGGGTCCTGCAGCTCCTCCAGCGTTTGCCTGCAGATCTGGAGCAGTTCTGGACGTGCGACTC
CCACTCGCGAGGCATTTCGACTCGGCGCGCGCGCGCCAGGCCCTTTCACGTGCGGCGAAGCCTTACCCGAGGTAGA
GGTCGACGAACTGAGCCTGGATGATCGGGTTACACTCAGCAACTATAGGTCCAGGTAATGCTGTTGCTGCTGCTGC
AGGCTCGGGTCGCGGCTGCGGTGGTCACTTCTCCGGTGATCGAGGATCCGCTGTTGATTTTGGTCTGCTCGAAGGTGTTT
TTCGCGATCTTGGTGGCCAGCGTGATACAGGCGTCTGGGGCAACTTGTGTTGGTGGTATCGAGAAGCCCATGCCGTGGA
CGAGACAGTGACCGATCCGCGTAGACGTTGTAGACGACGCCGAGGAGACTCATGTTCTTCGCGACGCGCTTGATTG
CGATCAGGCTGGGGATCAGGTTGGTACCGCTGGAGCCGTAGGCCGAAGAGGTTTCAGGGCGCGTGGCTTGGCCGAATG
ACGCTGATGTTGCTTGTCTCGTTGGCGTTGGACGAACTGAACATTCCAGCCGCTGCGGCCAGGCCACGCCGACCCC
GATGGCGATGATGATCAGCAGGATGATCATCTCGATGGAACGAAACGCCCTGCGTGGACGGGAAGTGCCTTGGCTAG
TGCTCATGGAGCGCTCCTTGGTGAGGTTTCACTGGTTTATGGATTGGATGAGCTGCTGTCCTGGTAGGCGCCGAGCAGG
ACCAGGATCATCAGCGCGCGGATGAAGATCAGGGCGAAGTTCTTACCAGCCCGCGCGGAGCTCGATCTGCTTGAAGCT
GGTCTCCTGCCAGCGCGGCTGAACCTGACAGCGCTCGGAGAAGCCTCCCGGTTGGCGAGGATGCACAGGTACTGGA
TGGCCTGCCGCTCGGGGAAATCGTGACCGCGCTGCGAAGGGCAACACCAAGTTCTGGCCAGGCCACGCCGTAGCGG
GACGCTTCCAAGCGCTGCTTACGCCAGGGCGGGAGATCTTGATCATGCTGGCCAGGCTGCTGATGGGGCGTATGCCGGC
GTTGAGCATGACCGCATGTTTACGAGGAAAGTGGTGGCTTGGAGCA'GCGGTAGATGGACAGGGCGGACGCGTCCGCT
CCAGCCAGACCCGCGCTTTCAGCGGTAGGTGCGCAACGTGACGATGACCACCACCGTGAGGGTGATGACCGCGACAGA
ACGTAGATACAGGTCCTGTGACGAAGCTGGCAATGGCGTTGAGCGTGGCGAGCGGGCGGTCAGGTGACTGGGTGGA
GAGCTTGGCAGGCTGGGGACCATGCGATAGGCCACGATGCACAGCAGGAACCATCATGGCGGACAGCGCGGAGGGGT
AGAGCAGCGCTGCCAGATGGTGGCGCGGATCCTGGCCTGGGCTCGACAGGAGAGCGGCATCGCCATGGCTGAAC
AGGTGCGGCTCATCTCGCGCGCTCGATCAACGCTCGCTCCTGGCGGGGAGATAGAGCGCCATGGCGGTGGCCAGACG
CTTGCGGTTGGACAGCCCCATCAGCGCTTCGCGACTGGCGATGGCCACCGGATGAACCGGATGCTGCCCTCATGAGCGA
AGATCTTGTGCACTCTGCCACAGCATCCTTCAACGGGACCCGTTTTTCGAGCAGGGTGGACATGCTTTCGTAGAACTGC
AGGCGTTCTTTCGCGCCGAACGTTGCTGTGTAAGGGCAACTGCAACTGCTCCAGAAGCCCCCATCAGCACGCTCC
TGGCAGTAGAAGCTGTGCTCGAGCAGATGCTCATGAAGTCAAGTGGCCCAATGAATCTCGACCATCTGCGGGTGA
CATGCCCTCGTTGATGCGGCGGATGGCGTGGGCGTGTGTTGATGCCCTGATGGTCTTGACCCAGTAGTTGCGTGCCT
CGGCTGGGCGCCCTTTCGCGAACACACGATGAAGCGAGGGTGGGCGAACCCTCGGCGACGATCGAGCGGCGGTTG
ACCCCGGAGCCACGGCAGGCTGGCAGCCAGGCCCTTTCAGCTGAACCTGGGAAACATCGGTCAAGCGTCGACCCGCTC
GACCAAGTCGGGCGCGAGTTGGTCTTGGTGGTCTTGAAGCGCACTTTCAGTGGGGGAGAGCTTGGGCGACAGGCTCT
GGTTGATCAGGCCGGTCAGAGGGCGGATCGAACAGCAAGCGGGTTCGACGCCAGGTCCTTACGGCGCTGGACAATG
CCGATCGCGCTGTTGGTGTGACGGGTGACCATAGGCGGTGCCGCTCATCGCACCGGAACCGGGCGACAGCGCGGAA
GAGGTCGCGTACTTCGCGCATCATCATGATGTCGGATCCAGGCGCATGGCGCTTGGCGATGCCCGCGCCAGGCTGGC
GTTCTGCGTCTGGGTGCGTGGCGTGTAGACAGTGGGCTGTTGATGCCCTTCGCCGGAATGCGGTATTCGGCGGA
TCCTCGATGGTCAGGATGTGCTTGGATCCGCCATGGAGCTTGTGAGGCGCTTCCAGGGTGACCTTCAAGGTGATGACTT
TCCTGACCCCGGTGGGGCGCGACAGGATGTTGATGCCGTAGGGCATACGATCATGCGATCGAACAGTGGCTTCTGCTCGG
GCAGGTAGCCGAGTGTCTCAGGCTGTGAGGCGGCTGCTGCTGATAGAGCAGTTCGAGGATCATAGGAACCCCGGCA
CGCGGGCGGCTGGCGATCCGGGCACTGAACAGGTTGAGCTTTCGACGAAGGTCTGGCTCATCCGCGCTCCTGGTCCAG
TTGCGGCTTGAACAGTGGCTCGGCCAGCTCGCACATGGATTGGTAGATGGTTCACAGAGTTCTGTCAGCTCCTGGCTGC
GGAAGTCTCGACGGTCTTACAGAGGCGCTCGACGCGGAAACGATCTTGTGCGGCTGCGCGCGGAGTACACGACGAA
TGACGTCAGTGGCGCGAGCTCAGTGGCTTACGAATGATCTTGACACGCGCTCTGATCTGGGTGGCGCGCTGGCC
ATCGCTATCGACAGGCGCTCCATGGCGACGCGCGGTAAGCTGGTGAATGGTCTGCAGGTGGTGAGGTTGAGCTGGT
ATCGGAAGCCACGGCGATCGAGACGGTCGATGAACGACAGAAGCTGTATGCTGTTCTGGTGGCTCTCGATAGATAGAGC
GTGCCGTGCGCAGCCAGAGCATGATTTCTGCAAGTGTCTGATACCTCCATTACACCCGTCGGCGGTGAGCAGTTG
GGTACCATGAGGGGTGCGCAAGCGCGCAATCTGAAGGTTGCTCATGGAGGTGCTTACTGAATGAACGGCTGCGGTAC
AGCCCCGGGCGAGCGCGCGGGAACCGAGCGCTTGGGCGGTAGAGAGGCTTGGGTGGGGCAACGCTGGAGAAGCCCA
CGGGCACGCGGTTGCCGTCTTGTGCTGAGCAGCTGTTCCAGCGAGATTGACTCGACGCGGTATTGCGACGCGAGC
TCCGCGCGCTGCGGCGCTCAACCTCGTACCCGCGCGGAAACAGACGTCGAGTCATCTTGGCGCGCGCGCTAGAT
GGTCCGACCACTGCGGGAACCGCTGCGTGATCGTGGCGCGCGCGCGGACCGGATGGCGAACGCTGCCCCCTGGG
CGCTGACGACCTGGCGGTTTGGCTTTCGCTTTCGAATGCGCTGCGCGGTGGCGAGGCGACCTTGGCTCGGTGAGG
ATGGCTGGGCTGGATCTCGCAAGTTCGCCAAGGTGCCAGCGCAGGCTGGCCAGACGCGAGGAGGAAGAGCACCGG
CACGGCATGCCGATCGGCTCAGTTGCTAGATTTCTCTGTGACAGTCCAGTCCAACTGGCTGCTTTCGAGCGTGGT
TCGAGGTTGGTATGCGGACACCGGACGCGGCGAGCCCTGGAAGGTGAGTCTGCGGGAGCGGGTCTGGCGCTGAA
GGTGAATTTCTTCAGGAAGGCAACCACTTGTGTTTCACTGCGAGCTTCCGCGCCAGGAGGGGCGAGTTGTCTCT
GGCTGATCGACAGCTTGGGCTCGACCCCTTGACGCTACAGTGGCTCGTCAAGCGCTGCAAGCTGCTCCGCCGCGAGT
AGCGGCTCATCACTGCCGATGGCAACCTTCAGATCGACCTTACGGGCGCGGTTGCGGTTGTGATGACGAAGGCGGG
GCGGTGCGGGAACAGGTGCTGGCTGGCGCTGTGAGGTGCGCTGCTGCTGTTGCGGTTACGGTGGTAGGTGGCGACCA
GGACCTGCGCTGCACTTGTGGATTGGAAGAGCCAGCCCTGGATCGACAGCGACAGTACCCCGTTGCCCTGCTACAG
GCGGTAGCATGCTTTCAGGTCAGGTATGAGCGTCAAGGCTTCTGCAATGACGCCAGGTCCAGCGCTGCTTGGCTG
GGCTTCTTCTCGGCGAGCTCGGCGAGGCGCTTGTCTTTCGAGGAGCGGCTTGGCTGGCGAGCTTCTCTGGTGGG
CATTCATAGGTAGTAGGCGGTAGCGACCCACCTACCACGCAACCGAGCAGGGCCACTGCCGTCCACTCCCTGGCGGAC
AAGCCGAAGGTGAGTTGCGGAGGCGGTAGTCGCGCGCAGGCGCTTGGCGCGAGCAGTTCTCGATGTCGAAGTCTCT
GACGGGGAATCGAACCTTTCGGGAACGAAGACTGTGCTTTCGAGCAGCGCGCGCTAGAGAGCTTCTGACCCGCT
CCCGGCGCTCGTCGAGGTTGGTACAGATCTGCCCGGGAACATCGCGCATCGAGCTGGCGACAGCGGTAGCGG
TCCTCGTCGACTTTCAGCAGGCGGAAGTCGCGCTGAACTGGCTGAAAGCGCGGAGGCGAGGTACATCCCTT
GACTGCGCTTGCAGTTTCGAACGAAGCGGCTGGATCAGCGTGGTGAATGGCGGATGGCAACGATGTCCAGATGCT
CTTCTTGGCCAGCTTGGCGCTTCTTCACTGACTGCGCTGGCTGGACAGCGGCGCGGAGAACAGGCGCTGACGAAC
TTGTTGCGGTGGTAGCTCAGGATCGAGAGCTGTGTCACGCGTGGCTGCGAGGTGAGGCTTCTCATGGTTTCAAGGCG
GACGGGGGTGATCAGCACACGATGACTCGCGCTTGGTATTGCGGTCAGCCCGCGGCAAGACCGAAGACCGCGT
CGCGGTGCGGACCTTGTGGTGTCTCGGTGGTCTGGTGAAGCGGAGAGTACCAAGGTCTCGCGCTGCGCAGACGT
ACCTTCTGGTGAACAGTTGTATGTGCTAGCTCGGAACTGGGCTTGGAGTCCCGCTGGTCTGATTTTCGAACGTCGG
CCGGGAGGTGATGTTGATGTTGATCTTACGAGCATCTCGCGCTTTCATCAGAACGCGCAGCAGGCTCATGTTGTAGC
CGTGGTCAGCGCGCGGGATCAGCGAGGTGGTACTGCCGACCTGGGCGAGCTGGAGATCTGGCTGGAGGCGAGTAG

Figure 2

FIGURE 2K

CTGCTGTAGCGGCCGATCTGGATCGGCGCCGACTGGAGGTTGAGCGTGGTCACGGACGGGGATCGGACGGTCGAGACGCG
GCCCTCTCGGGCCAGCGCTCGACCATGGCCCTTGGATCTCTGCCAGGCGCTGTTGGCGGTATCCAGGATGCTCACGGAGC
CGGAGATGCGGCTTATGATCGATCCCGGCATGGTGTTCTTCAGGCCGATGCCCCATCTGTTGTTGAGCGCACTGTAGACAC
AGGTTCCAGTCGATCCCACTGATCTCTTGTGCTGAGGCGACCGAGACAGCTGAGCTTCAGGACGACCTGCTTGGT
GATGCTCTCGTTCTCTCGGTTGACCAACTGCTGGACACGGTTGAGGACTTCTGGACGGTCGGTGACGGTCAGGGTGCCCG
TGGCACGCGACAGCGACATCGCTCCCATGCTCGGCCGTCAGCATCAGTTGATGCTGTTCTCGATGTCGCTGAGGATCGAT
GTCTTCAGCTCCGAGCTGGTGGTCTGCTGCTCGCGGAGTCGCCGCTGATGGAGCTGCCATTCTGTCGGTGGATCC
GGAGCCGTCGCCGCTGATGCCGCGCGCGCTGCTCATACCGGAACGACGCTGAGATCCACGTTGACGTGTCGAGG
CGTACATCGGGAAGGTCGAGTGTCCAGGTAGTAGAACTCGAACCTTTTCTCGTTGGGTTGTAGCGCCAGGACACGCCG
GCTCGGGCGGCGATGAGATCGAGGAACCCCGCTGATTGGCTTCATTGATGTTGTACAGAGACCGCGGACCGTAGGA
CCCGAAGCTCGAACCATGTCTGCCCGCGGACCGACGACATGCCGTTGGCAGCGGAGGAAACAGCATGGTGGCCCA
TGTCCTGGCGCCTTGGATGGGCGGCGGGCGTTGCTCGCGCGTGTGAGGTTGACCGGCGAAGCGCGGGTTGTCAGC
GCGTCGGGCGTGATACTGACCGCCATGTGGCATTGGTTGATGACTTCTGGCGCGCTCTGCGAGCGACGCTGCGCCTGC
AGGGCGCCACGTCAGGTCAGTCACTGGACAAAGGTGTGCGAAACGCTTAGGGGTTTCGTGCTGACCCAGGGTTTGTGCG
AGAACACCAACGATCGCGCGCATCCGCTCGGCTGTGCGCACTTGGCTGCGATCTGAGACCGGGAATCTGCCGTAGCT
TCGACACAGATCCGCCGACTCGTTCACCCGCGTGACCGAGCAACTGGCCAGCGCGGGATCAGGAGGCAAGCCAAAGAGG
TTCTTCAGGGGGGCGACGATCAGTTCTTGGCCTCTTGATGATGATCAGGAGATCGGGCGGCTGGCGTTTACCAGGAAG
GGCCGTTCCGCGAGCTCATACAGGGGGAACAGCTCGGATACCGCGTCTCTGGAAGGAGCGTGAAGGTCAGTGGAGCCCTC
CATCGGATAGTTGAGATCTCTCGGCTCCGACGGACGCTCGAGCTGCGGCTTGGCCGAGCTTCAAGGTGTCGCGCA
GGGTTGATCCCGTCTCGGCTGACAGGCGCTGTGCCAGCGGCGGCTCGGTTCCGGCGGCGTGAGCACCTTGACCGGTGGG
-GCGGAAGCCACGGTGGGTGGATGGGCGGCGTGGACTTACCGCGGAGGCGCGAGGCTTGGTGGTGCCGAAGCCGCGGG
AGGAAAGAGGGGTGGCCATCTTGGCCGCTGGCGCTGGGGGAGCAGCTGGTGAGTTGAGGCTCTTTGGCCGCGAGCGT
GGCTGGTTGTGATCGGCGCGGCTGTGCGACCGAGGCGACAGCATGGGCGACTCAGCGCTGCTGACTTTTCTCGTGAG
GAGGGTTGCGCTACCGCGCGGCGAGTTGGGGCAGCGGGCTTCGCATACAGTTGCTGGACCGGTTTTCGGCTCGGCGCGG
GGGAAGTTGATAGCCCGGGGCGAGCAGCAAGCAGCTGCGCGCGACCTCTGCGACCTTAACTGCCAGGCTGGGCGG
AGAGGACCTGGAGGGTTTGGCAGGGTCATCGGCGCGAGCTTGACTGAGCTCCGCGAGCGGCGGGGTGAGGATG
TTACATGACCGGCGCTCGCCGGGCACAGCGAGTAACCCGAGCGGCTCATCAGTATGCTGAGCTGCGGAGCGGCGGGGT
GTTTCATGCTCGACGGGATGTTACGTCGATGATCTGGGCCATCAGGTCGGGTTGACCGGCATCAGGCTGGGTGCTGACCA
GGGTGATGGCGCCATAGCGCACTAGGGCTCTTTCTCGGCTGCACACCGTTTGATACAGGTCGCGGAGAAAGCAGGTC
GGATTGCGGCTGCTGTGCAAGGCTGCGCTGCGCGGCTGACGCTTGTGGGTGCTGAGCTGGCGAGCAGGCAAGCAGCT
GAGAGTGGTCACTGCTGGTCATTGAGGCCCTTTGATGATGGGGCACACAGTGGCAGGCGCATCCGCGACGATCAGC
AGGAATACTCGCAACAGGGGAGGGGATTTTCATGATCATAGGCAGAGGGTGAGCGAGGCGCGTAAGGGAACCGGAT
TAGAGCGATGAAAGTTTTCATGTTACGCTAACAGCGAGCTGAGTGGGAATCGTTGTTAGAAATGGATAAATAAGAT
TCTAATCGTTTTAGAAATGTTTTGTTGAACCTAGGTCAGACAGTACGGGGTCTAATTGAGCTTTAAATACTATTGGGCTA
AATTTCTCTGGGAGCTTTCTTCTTGTGTTTGTGCTTGGAAAGGCTCCGATTTTCAAGGCTGGCATCAGTTAACCCCT
TTGGTAGTGCTGGATTCACTACTACTCTGTTTCTCTTCTCGGTAAGGCTGTTGTAAGACTTCGCTTTAAATATACGGA
AAAGAGTTCTGGGTACAGGTTTTTTCTCGAAACCCCTGCAAAACAGGATGTATGAGCTTTTATTGCTTTGTTGCT
ATTTGTTTTCAATTCCCTTGGGGATGATTTTTTTATTCTATAAATACGGAAAGGCTCGTAGGATTTAACGAGGCTTGC
TTGGTTTTAAATGGACATTACAGGTTGTTTGTCTTTTCTGAAGGCTTCGTCATATCGCCCCGGTAACCTGCCATTA
CCAGTGGCAACCCAGGATGCTATAGGAGTGCGGCTTCCGCTGGCAATGCAATACCAAGCCTAACCTGGCGCGCGCA
CCAGCAGCTAGTGTGTTTCAACTTTTACAAAGAATGTTTTCCAGTCCCGGTTTCGTAGATATCTTGAATCTTGTGATG
GCTGGCGGCGTCAATAGCTTTGCGCTACAACCTCAAAGCCTTGTCTAAATTTCTCAAGGCTCTTCGCCATCATCTGCTGT
CTAGAGAATCAAAGGCTTTGGCAATGGCTTGCCATCTTTAAGGCTAAGTTTTTATTAAAGCATCTTGTGTTTTTCA
AAGCAATGATGTTGCTTCGCGCAACTCTGATATTTTTCCCCCTGGCGCTTCGGCCAGTGGTGCGGCATCTCTGATG
TCGTGCGCCAAATTTCTCAGTTACTTCTTATAAAGTCGGGAGTAATTTAACAGCATCTGTGCGAACGCTGTGCTCGG
CCGCTTTCGCTTCTGCTTCTGCTCTGCGCTGCTGCTGTTCCGCGCTTTCAGCTTATCTAAGCGGCTCTGTAGACTGGCTATC
TCATCCCAAGTAGCTACCACTGTGATGGCTTTCTGCTTTCTAGAGGGGTTGGGCGATCTTGTATAAGAGTATTGTT
TGTGCTCTGCACTGCTTTTTATCAGCAATTGCTTGCGTAAACGACTTGTGGGACGTTCAAGCCATTGCTCTGGCTTG
GACTGACCAATGGAGGCAAAGCACTTCCGCTTGATCAAGCTCGGATTGACTTTTATCTCGAGATTGCGCTTCTGCTGTT
TTGTAGTCCCGCTCCAGTGTAGAGAGGTCCTCTTTTCTTTAGCGCTTTTGTGTTGAGTCTCTGTCTTAGCACTTT
TACCGCTTCAGAGATGAATCTGTTGTAGTAGGCAACCGCGGCTCCAGTTCATCTACAAGCTGTGGAAGGGGT
ACTTTCTTTCTTCTTCCAAAGAGTCTCTTCCAGATCTGACACAGGCTCATAGTAGGAGAGGCTGACTGCTTTTCG
TAGTCAATGAATGGGCTTTGGTAAAGGTCTTTTGTATCTTCAAGATCAGATTCATCTTGTGCGAGGACTGCCGACAT
GACATCGACTCGCTCTTGCTTTGTGGCCTCCCTTTTCTTCTGAGTATCTGCTTTAGCAAGCATCGACTGATCAGATA
CTTTCTGTGATCTCTCTTTCTTTTCTTCTTCTTCTGTTCAACAAATGGATCGGATGGATAATAAATGTTTTCTGTAGA
TATATGTCTCTTTGAATCAAGTCGCGAAACATATTCGCGCAACGCGCGCTCATCGATAGGGGGAACCTCGTATGCGTA
TTGATCGCTCGGACCTTGTGCGACATAAACCATGGAACAGGTAATCTGTTGTCATTGGACATTTAGACTTCTCCATTGG
TGAGTGTGGTAGGCAAAACAGTGGCATTGAGTTCAGTATAGAAATGAAACCTAGGCTCAATGTGTGGGCAATGCACTAAGA
AGGACAGCTCAATGCGCTGTGATATAGTGGTAAAGACGGTGGTCTTAACTCGATATCAGGAATGGGATGGCTGAGC
CCACGGGATATTCCTAAGAGTCATGTATCTCTAATAACAGAGTTTCATGCGCTCAAGATTACAGCTGGGTATGTGATGCTC
TCTGCAAAATCAGCTACTTGTATCGCAACACCATTAAGGGTGAAGATAGGATAGCAAGCGAAGAGACAGTTTTGTGC
GTTCCATGTTAAACGGGAGATACAGCTCGAAGACCTAAGCCATTAAACGCTGTGACCATCAGTTGGGGCCATTAGGA
TCTAGCACAGCAACCATGGTGCATGGGGCTATTCTCGATGTATACCGTTTGGTAGAGTTCGGAGAGGACAGGCT
GGATTGCGGCTGCTGTGCAAGGCTGGCCTGGCCGGCTCGGCTTGTGGGTGCTGACAGTGGCGGACGACGACGCGT
GAGAGTGGTCAAGCTGCTGTGATTGAGGCGCCCTTTGATGATAGGCACACAGTGGCAGGCCATTCGCCATGCTTCAGCA
GGAATACTCCGCGGAGGATCAGGCTTTTTATCCAGTTAGGCTTGGGCTAAAGCGCGCGGTTTCCGTCGACCTGCT
CGGGCTCGATGAGGCTCTGGGCAACCTGGCGTTTTCGCTTGGAGAGTGGCAAGTGTCTGTCATGCTGCGCTCGAC
AGTGGGATTTTACACAGCATGCGGAGCTGGCGGTTTCGGTACGCGCGGCTTTCGGCTTGTCTGCTGACCGGAGT
CCAGGGCAGGCGGAGGAAAAACAGTAGTTTCGCGCGAGTGAAGTTGTTGCCGCTCCCTGCGCGCGGCTGATGTCAGATGA
ACACTCGGAGCTCGGAGCTCTGCTGGAAGCGATCTATCGCTTCTCGGCTTGGTGAGCGAGTCAATGCCCAACGAGCTG
ACGACGCGGCTGTCGGGCTGCTCGCAGAGTTCTTTCAGCGAGCACAGGTCGGCTGGAAGTGGTGGAGGATGCTGCCCAACGAGCTG

FIGURE 2L

GTCTCTGCGTCGAGCTCGCTCAACAGGTCCATCGCCACGCGAACTTTACCGTTTCCAGGTAACGCCGAGCGCGCGA
GTCGCGCGAAGACCGGTTCCTCGAGGCGCAGCAGCTCGTATTGCTGGCGTTCCTCGGTGGAGAGGGCCACCTTCAGC
AACTGCGCGTCTTCCCTTGAGGCTGGGCGAGCAGCTTTGCGCTGCGCAGCATCCAGTACCCAGCTCCGCCCGCAG
ACTTGGCGGAACTCCGGGTGCGCGGCGAAACGGTCGAGAACTTTTACGCGCAGTTGGCCGATGGGGTGGCTGAGA
GGCGCAGCAGGGTGTGAGCTCTGTCTCGCGGTTGAGCACCAGCGTGGCGGTAAAGCAGGTAGCGGTTCCGCACTTGGGCG
GCAATGTGAAACCGTCCGCGGTGCAATTGCGCGCTCGGTTCTTTCATCCGCTGCGCCTCGTGCATGACCATCACGGCGAA
GCGCGAAGCGTTGGCGACGAAGGGGCTCAACTGCTCGTAGTTGACGAGGATCCACTGCGCTCTGGGGTGTCTGCTGGA
TGGCCACGGTGGCGGAGGGATAGACCTCTGGATCTCCGCTGCCAATTGATCAGCAGGGTAGCCAGGTGATGACCAAGG
ATTGGTCTGCCCCGCGCGGATCGAAGCGCGATGACCGCTGGCGGGTCTTGCCCAATCCCATGTCTGCTCGGCAATAA
GGGCTGTGTTCTGTCAGCAGATGAGCGATGCGCTCAGGCTGGTGGGCCAGTAGAGAGTAGCCCTGAAGCGCGCTTCGA
TATCAGCCGAGCTGTACTCGGTGCGCTCGATCTCCGCGACCGCGCGAGGTAGATGTGAGGGTGAATCTGCTCTCCAGG
GATGGCGCTGCGGGTCTGCTGTGGACCGCGATGCTATGCTGGGCGAGCTCGGTGACGCGCGAAGTCTCGCTCGCT
GAGCAGCTCTGACAGATATCCAGAATCTCGAATTGTTCTCGCGAGGCGCAAGTCCAGAATCAGGTTGCTTCGAGCA
GCTCGGGAGTGGTGTGATGCGCCAGGAATCGATGGGCGCAGGAATACCCCGCATCCGCGGAGCCACCGCCACAACA
CCAGGATGGAATCGCTGACAGTAGGACGCGCGCTCGGCGAGGGCGCGATGCGCAGCAGCATCCCCAGGTGAAGGC
CTGTGCGTCAAGCGAGGCTGCGCGGATTCGACCAAGGAGCTGAATCCCGCAGGAGTGGAGTGAAGGCTCCAGTGA
GCTCGGCGAGAGCGGTTGTACAGAGTGTCCAGTTGAGGCGAGCAGCTTGCTTTGGGGACCCGCCAGTAGCGGTGCAACGA
TGCTCCCGGGCTGTATGAAGAAGCGGACAGCTCCAGCAGCGCGGATGTGCTGCTGCGATGTAGGGGATCTTCAAGCC
GAAGTCGTAGCGCTCGAAGATAGCCGCTGACAGAGCGTGAAGGTGTGCGGAGGTACGATTAACAGGCGTGGACGGT
TCATCGTCTTCAACTGAGTGAAGCAGCACTTCGCGGTGGCTTCGCGAGGCGCGCTTCAGCGACCATTTGGCGAAGC
GCGTCATGTGGAGGTCATGTGCGGCGAGTTCAGGTTCTGCTGCTATGGTCCAGAGCGAATTCGTTCCATGAACAGCGG
TATGCTTGGCGCGCAGGEGAGTGGTGGCGCAGGGGCGATTATCGAATTCGCGCAGTTGACCGTGGAAAGGTGCTTGGG
ACTTCAGGGCTTACCTATAGGCTCGCATGCTCAAGCTGAGCTCAATTCGCGAGGCGCGCTTCAGCGACCATTTGGCGAAGC
GTACTGAGAGCAGTCGCGAGTCTCTCGGTAGGCGCGCTGGTCTGAGCTATCGATGCAATGTGCCAGAGCGATATGTTG
GCATATAGGTGCGCGAGGACAGCGAGCTCGACGATTCGAGTTCATTAGGGCAATCGATAAGACCGACAGACAGCATTC
CTCCAGGTGGTGTACTGCTCGGTGTGACACAGATCGGGCAGGCTGCTTCGACTGCGGCACGAGCTCCCTCTGGACCT
GGTGGTAGTCTGTCACAGGTTGCCAGCTTGTGCGTAGGTACAATGACCCAGATACGCTTTCAGGGCGCTGCGCGCTGCG
CGAACAACTGTGAGTTCCTCGGTGCGAGGCGGTATAGAATCCTTGATGTCTTCAGCCGTGAGTATTGCAAGCCA
TGCGGATCTTCCCAATCCGACGCGGCGAGCGGTTGAAGAGACGGTTGTAGTGGGCTTGGCGTTGTGAATTCGCAAGGGA
TCGCAATGTAGGCGCGATAGAATGCTGTCCATTGAGCGCGAGTGTGCGCGCGGTTACCTAGCGATGCGCGCAGTAGG
AGAGCTAGTCCGCAATGTGCGGCTTCCGCTTTCGATGGAAGAGAGCTCCTCCGCGACGCGCGGAGCTGACGTGACT
AAGAAATCGAATCGAGGCGCGCGCGCTGAATAGCCAGGCGAAAAAGTTGATGACCGAAATTCCTTCGTCGCGACCT
GCCCAAGAAATACGGCCATATCGAGGTACAAGTTCGATGCGATGGCGCGCGCTGCGGTGCTTACTGTCAAGCTGGGGA
AAGGGCAAGGAAGGTAATGGCGGATGCTCTCGCGTTGTGCTGGTGTGCTGAGGCGAGCTCCCTCTCGCAATC
CACAGTAGGCGCGCTGAAGACCGGCTGAGCCTTCAGCAATTCAGCGCGCGCACCTGGTGCCTTGAGCGAGCAGTGT
AGAGCTGGGCGGAGAGGCGACCGAACCGAGCGGATGGCGAACAGCAGGCTTGGTCCAGGACCTGTCTGAAGGTTGCAAG
ACCGCTCCGATCTGCTTCAAGTAGAACTGGTGCCTTGGCGCGTAGTAGCGCTTGTGCTGTGGCTCGGATACAGAGGCG
ATGTTGAAGTCCGAGATCGTCCCTGGCCATCTGTGCTGCTAGGCTGCTGCGCTCGCTTGTGTATGAGCTGAGTACAGCAA
CTGCGTCTGCAACTTTTGAAGGCGCTGTAGCAGCTTCAAGGCGCTGCTCGAGACGCGGTGCTTGGCGGATACGATGG
GCAGGCTTGCATCCACGACCAATAGTTGACCGTCAATACCTGGAGTAGCGGCTTGCACGCATGATGGTGAACG
CACACGTTGTTCAGCAGGACTGTGATGACACAGCAGCTGGTGTAGATGTGCTGCGCAGGTAGTGAAGGATGAATCTGG
TTCGAGCTGTCTTCGAGGGAAGAAATGTGTTGCCAGGTGTGCGCAGCGCGAGCTGCCCTCGCTGTAGCGGGCCAGG
GATGGATCTGTAGCGGCGCTAGGGGGGCGTGTGCGAGAGGTGCCACGCGATGCTCCTGTGTAAGCAAGCAGCACT
TTGGCGCGTGGTTTACGCTGAATTCATGTGAAATGGTCACTCAAGCGTGTGGATTTCAAGGTTCTACAAGCTAGCTG
ACTGAAGGCTTATTGGGATGGATCAGTAAATCTCTTGTAGTCAAGCTCAAGTCTCCTAGTGACGAACTCGTCAAG
GATAGCTTTACTTAAAGAAATAAGATGATTTGCTCTGCTGGGTTGTCCAGTTGGCACCACTATGCTTCCCTCGTGAA
ATAGATTGTTTTCGACGCTTAAAGTAAGCGAGCAGCTAGCTAAGTGCCTTTTGTCTCATCTAATTTAAACCGGACGC
CAACCGAGCTCTTACCGGGCAGGACTATTTGTTGCTCTGAGCTGTCTCGATTAGTGTGTGGCTGATTGGGAAAGAGA
GTATTTGTCAGAAATGTTTGTACAAATTTTCCATCCCGGCTTGCCCTAGCTCCAGGTTTGTAAATTTTAAAGTAGC
CATTTCTTTGAGGCTGAATCAATCGCGAGAACCAATAGAAAAATCAAAAGCTCGCTCTGTGATTTCGGCAGATATA
TTCTCTTTTCGATAAACCTCTTGTATAGCTAGCAGCAATTTGGGCTGTACCATTTTGCCTGCACTGCGTGACTAACT
ATAGTGGCTATAGCCCGCGCGGTTGTGCAAGTCAAGCGCGCGCTCTTGGTGGGTAGCTAACTTGCACCTGTGTTGT
CGATAGCGCTGCAACCCAGAAAGTCAAGCGCGCTTTGCGGCTTCTTGGGTTGAACAGGGGCGAGCCGCACTTGCC
GCAAGGAAGTCTGCTGGTGGCGCGGAGGATAGTGGCGGTGGAGCTTGGCGGTGGTTCTGCTCTGTGCGCTGGGAGCG
CGCGAGTGTCTGAGCTCTGTCTGTAGTAGCTGGTAGAATGCTGGACACGTTTATGTAGGTGCTCTGCTGTTGGCG
ATCCGGTCAAGGCGACCTCGAGGTTGGCGGTGAAGTCAAGGCTGAGGAAGCTGAAGTTGTGCTCCAGGAGCGCGATGGT
GGCTCTCCAGCTCGCGGGGATCAGGCTTCGCCCTTCACTGCACTGCGCTTGTGCTGATGATGTTCTTCAAGATCG
AGGCGTAGGAGGATGGCCGACCAATCCGCGCGGCTTCTCTGCGCCAGGCTGGCGTGGTGAATCGCTTGGGAGGG
GTGGTTTTCTTCTCAGGACCTCGCGCTGTAGACCTTGAATGCTGCGGTGGCTCCAGCGCGGGATCGGGATGGGGTT
TGAAGGCGTTTCTTCTGCTCTCGCGCTCATCACCTGACGAGTTTTCTCCAGCCAGGACGTTCAACAGCTTCC
CTTTGGCGCGAAGCGCAGCGGCTTTTGTGCGGGCGAGCGCCAGGAGGGTATGGTTCTCACTGCGTACACGGCAGCT
TCGATCTGACTGGCAAGCGCGCGGACTCGAATGAGCTGGTACAGCGCTGCTCATCAGAGTTTACCGCGCGAGCGG
CATCAGTGGTAGGGGTGATGGCGGGTGGCTTCTTGGCGCTTGGTCCGCTTGAACATCCGCTGTTGCTCAACAC
ACTTCAGCCCAAGGCTTTGGCGACAGCAGGATATCCGGCATCGAGTCTTTCAGATATTTGGGGTTGTCGCTCGGTTG
TAGGTGATGAGCCCTGTTTATACAGCGCTGGCGCACTTTCATGCTTGTGCGGTGACCACTTCAAGCGATTCCCGGC
GGCCATCTGGAGCGAGGAGAGTGAACGGTGGCGAGGCGCATCGGTTTCTTGTGATCAATGCAAGGCTTCGACGATGA
CATTAGCTATAGCCCCACAGTTTCTGCGAGTTGACGATCTTGGACATACGGGAATCTCTGCTGGCAACACGGGCACT
GGATCCCATTCGCGCTCCAGGTACGGCGCTGCTGGCGAAACGAAGGTGAGTTCCACCCCGAAGTGTGTTGCTGTG
GAAGGCGCGGATCTTCCGCTCTCGCAGGACCAAGGATACCGCGCAGGACTGACGCGCGCGCGGCTGGTGGCTGAC
CCATCAGCGCGCAACTCTGGCGTGACAGATACCCCAAGGCGATCGATGACGCGACGGCATTCCTGCGAGGCGGAC
TTCGGGAGGTCCAGGCGACGCGCGAGCTGAGTTGGCGGTGATGATGACTTTGTGATTCTTTGAAGGCAACGCGCTT

FIGURE 2M

TTGATGTTCTTGATCCCGGCAGCTTGGAGGATGTGCCAGCCAAATGCTTTGCGCTTTCGCGATCCGGGTCGATGCGACGTAAGA
TTTTCACGGGTTTTCGCGCGCAGCTCCTTCAGCCGTCGACGGTTTTTTCCTTGCCCCGAGAGGATCTGGTAGTGCGGTTTG
AAATCCTGGCCCCACGCCAGCTGAGCATCTCGCGGATCTCCGGCTGAACGGGTAGGTCTCTGATGTGCGCGGCTGTCGC
CTGCAGCTGCCAGTTCAGTCGGGACGAATCAGGGGAGAGGAAGCACTCAGCTTTTTCAGTTGCCCCGGGCTTCAATGA
TTACGATGTGCATCGCGGAATCTCTCGGAACAGTCGGTCGAACCTGTCGCGATCAGAGAGCGGGGCAACCAATGCGAAATA
GACGGTGGCTGTTGTTCAGTTTTCGGGGCAACTTTCGCGCATCGCGGAAAAAATATCCGTTGCCGATCAAATCGGACACA
ACTTACCGATTCTCAAATGCGGTAATCTTGTCGGTGGCTGTCGAACACTCTACGGTAAATCGTGGCAACTGGGTGCGT
CGCTCGCATAGATTACAGACAAGAAATAGCATGCTCTCGGAAGGCATCAATCAGCCCAATCATCCGATTGGTCTGAGG
AAGGAAATTGGTATGGATCAAAGCCTTTGCACATGCATGCCAACGCCAATCGTCAACCCCAAGGAGCTGCGACTGTGCCA
CATGTTGATCGGTGAAGACTTTCCCGATAACATTTGATCGCAGCGGCACCATTTGGTTAGCTATGACGGCAGCGCTTGGTGGG
TCGATCGCGATGAGCCGCGCAGGAGCAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTCACTACGCTGCTGCTGGTC
GGATAGGAAAGCGCATCAGGTACACCCAGGCCCTCTAACTCGTGGGGCAAGGCCGTGGGTGAGCCACGACTTACTTT
GCCACAATGCGAAATGGAACGGTGGCTGTTGCGAGTCTTCGGGAGGAGCTGAAGGGTAAAGTCATCAGAAGTCGAGCTG
TGTTGCTGCTCTCTGGGCAACCCCTTCAACACGAGCGCCGCATCGCAGGATGAGTGAAGTGGGCTGCTCAGACCCCGTAG
CCCTTTCCGCGCGCTCAGACCAATAGCCCTACACTCCGCTGGGGATCGTTGAAGTGTCGTCTCGCATGCTGTCGCTG
TATCGGTGCTCTCTTTGACACACTGCACCTCGTACTGGTGCAGATTTCCGATTCCGAAGGCCCTTGGCACCCAGAGACTT
ACGAACGGGCGGCCAAAATTAAGGATGCTGGTCTATCAGCTTTTGCCGCATAACTGTTTATTGCGGTTGCTCCAGTCGC
TGGAACTGTTGTGCGAGGCTGCGGGCGGCAAGAGAGGTTCTTGCCACCTGCAACTTCACGAGTGGGCGGCAAAAATTAAG
ACCGCGCTTTTACCGTCGCGTCGGCAATGCTTGGCCATCTCGATGAAGATGCGCGATCTTCTCTCCATAAGCGCC
ATGCACTCTAGAACTGATATCTCGTCGGCGCTGAGGAGCTGCTGGTCAACGGCTCATCGCGATCCGGAATGAGCAATG
TGAGCTCTGTTGGTCACTCTTGTCTGTTTCATGGCGGAGTTTCTGCGGGGGCTTGTGCAAGCTGCGGTTGGCAAAAAGT
GGTACTTGGCACCCGAGCTTACGAATGGGCGGCAAAAATTAAGCCGCTGCTACATGGGAGTGTGCTGCTGCGAAGGC
GTCCTGCTAGTCTTGTCGATGTCGATCTTGCTGAGCAGGTTTGCTGCGGAGCTTCGAGAGTGTGATGCTGGGTTG
GCAACAGGGTGACCTCGGCCAGGCGATGCGGAAGGATGGAATGCGCGAGGCTTCGACCTTCAGCGCTGGACTTCTTCG
CCTGACTCTTTGTCGGCGCAGCGGTGCATAATGGCCCTGCTTCGACCTTGACGCGCATACCTTTGCGAACAGTTCCGCG
GAAGCTCTGGCATCTCGATGCGCAACTCGACGCTGCACGAAGCCGCGCGATCTTGTAGCAACTTGGCCATCGG
GAATCGAGTTGTGCAACATCAGTTTCACTGCGCAGCACTGCGCGGGGATGTTGCGGTTGGGAGAGTCTGCGTGCTCG
GGCGCGAGGCAATGTTGGCTTCCAGAAAGACGGGGTAGCCATGATGAATCTCCAGTCAGTGGTGGTGGTCTGCTGCT
CGAGTCGGCGCAATGAATCGGCTCGGCTGGCCCATCTTTCCGCGCATCGGCGGCTGCTGCCGATGAGGTGGAATC
AGGCTGATCTGATGTTGAGGGTGATGCGTGCAGCTCCATCTGCTAGAGTTTTCGAGAGTTTTCGAGCAGTATTCGGGTTCTG
CTGGCGCAGCATTTCGTTCCACACCCGACGCCCATCCGCGGGATGCAATGTGCTGCGACGCGAAGAGGTTGTCGGG
CGGATGCGCTGCTCGATGAGTCGGGTGGGCAAGTGAAGGTTGGGCGACGAACCTGGGCGAGCAGTCCCTGGGCA
AGTTCAAGCAATCCCTGCTCATGGCTTCGCACTGTTCCGCACTGGGCGAGCTCCCTATTACCTTTAAAAGGCTTTAA
AAGGCTCTTTAGTAGGCACTGTTGTCAGCGCTTGAAGGCAAGTTGTTGCACTGGGCGAAGGCATTTGTTCCAGAG
CGCTTGTCTGGTCTTATCGAGCTGCCGATTCGATCAGTCTCGAGTTCGACACGGTCAGCGTCATCATCGGCATCAC
GCTCAGAGCGGCCCGGGTGATCGCGGAGCGGAAGTTCAGCGCTCGAGTGCCCTCAGGATGCTCTGCGGGATCTCGCG
AACTCTCTGATACCTTTCGCGCGGCTTCGCGCGAGCATTTGTTGCGGCGAAGTCGTGCGAGTGGCGCCGAGAACTG
GTAGCTCTGGGCCAGACCAACAGGCTTCGAGCAGCCGACGCACTTCTGTCAGGCAGACTCCATGTCGCGGCGACGA
CCAGCCCGCAGTGTCTGGGCCAGCAGAGATTCAGCGGCGAGTTCGTCGTAGTTGGTCAAGAGATAGACTGCTTGAAGCCG
AGAGGTTGGAGATGAACACGGGACCTTGACCGGTGTAAAGCGACAGGTTCTCGGAGATATCGAGGTCGCGGCGAGCTT
GGCATGACGCTCATTCAAGCTGCTGGTCAGTGCGCTGAGTCGCGCTGGCAGCTTCGACTTCTGATCTGAACCA
GCCACAGTCGAGGTACGGATCGTCTGTGCTGCCCGCGGTGCATGCGATTACAGTATGCGCAGAAGCCGAGAGGCGC
AGGATGCTGTGTTGGCATCCGATCTTCGCGCGCGGTCAGATGCGGGCGCGGTGATAGGTGTGAAGGCTCAGAGCGAC
ATCGTTCGCGAGAGAGCGCATCGAAGCGCATGGGTTTCAAGCATGGTGCTGTTCTTTATAGGGAAGGGGTGCGCC
ACATCGTGCTGATCGGGTAGGAGCGCAGTAGGAAATATAGTGTTCGCGCTGGGGAATTTCCATGGGCATCTCGAG
CAAGCGACCCAGTTTCGCGCGCGCGCTCGAAGTCAGTTCTTAATGGAGTGGTTCCAGTGGAAACCCCAACTCGCCGTCG
CGCCGAAGTCAGTTCTGTTGGAGTGGTTCCAGTGGAAACCATCTTCGCGCTGGCGCGCACTTGGGTATCTCTACCGT
TTGGTAACTCTGGTTGGCAGTTCTAGTAGTTGGCTTTGTTGATCAAAGCTCCCATACGTTTTCGCCCATTTCAAAG
GATACATCTCAACCGTAATGGTCAACACGGTTACTGAATACCCGAACAGCATCAGCATTAACACAGCCTGAGTTCTCGG
TCGTTCCACGCTATCTGTAAAGTGACAGAAAGACGAGAACGACCAAAAAGAAAGGGGCCATGCTCAAGAACATTGTTG
AGTTCTCTGTAGTGTGATGCTCGCTTGGCGAAGTGTTCGCGCCGAAGAACCCATCTTCGCGCGCATACGAAGTTCG
GATCCCTGCGCAGAGTATTCAGTGGAATCACCAAGCTGCGCGCGCGCGAGGTCGGCTCTGGTGAGTATGATGAG
TGGAATCACCCAGCTCGCGCGCGCGGAGGTACAGACCTGGGAGAGTGATGCGGGCCGCATCACCAAAAACACGACG
CCCGACCCGAGATTTCTCATGAATCATCAACCCGCTCGATGGCAGCGCTCGATCCGTGCTTGGCGGCAATGAGGTCTTGA
AGCCGCTGAGGACTCTGCTCCCTGTCGGGTATCCCCCGGGGCTGGAACCTGAGGCTGTGAGGCGCTATGGGGCTTGA
TGGTGGCGGAGCACGAATGGGGCGTTTCGTTGCGGGGATGGTTGCTCAGCGGTTCTTCGGGAGCCAGAAGCGTTGA
ACTTGCCCTGGAATGACGCTGGATGAGCGTCATGAGGTAGCCGAACGGATTGCTGACTGTTCCGCGCGCACAAACGGTGC
TGCCACTGCTCGATGATGGGTAGCCGAAGATCTCCGGAACCCGCTGAGCGCGCTTACTGCTGGATCCGTTGCTCGGC
CTCAGTGCCTGAGGAGATCTCCGCACTGTCGATGGGCTCTTCGCGTCTCGCGTGTGACTGGCTTTTACATACG
AATCTGTTTGTATGTATACGTACTATAGATTGCGCATCCGAACAAAGGGCGATTGCGCGGATCTGTGCGGGATCG
GTATCCGAATCAGCGGTAGAGCGCATGCTTGGCGTGAATTCAGTGAAGTTCGGCATCCGAACCTCGGGTGGAAATGAGG
TAAATCCGACAGAGTTCGGATCGGAATCTGCGGCTGGATCCGCTTCGAGCTGACTATCGATAGCCCATTGCTGGTGGT
TCAAGCGTCTCTCAGCTTCTCAGATGGGTAGGGAGGCGCGCACCACTCGGATCATCTCGTAACCTCCGCACTCCCGCAGATG
ATCTGCCGATTTTCGCTATGGCTCGGTTGCGGTGACCGGTGGATTGGCTCAGCAACTGCATGTAGTCGGTGTCCGACTC
CAAGGCTTCGGCTGGAGAGACAGGCTCTGTCGTGAAGATGTAAACAGTTGCCCTGGACGTTGCTCGTTGAGGTGCTTGGCA
GGGCTCGCGCGAGGCTGAGCCAGCGGTCAGACGAAGGACCGTGAAGTGCCTTGGGAGTGGTTTCGCGCGAGGCGACTTGG
CCGCGCTGCATACCGAGATAGGGGCGAGTTGCTGTAATGTCGGGAACCGGTGAGTCCGCTGCTGTTGATGAGCAAGCG
GAAAATTTGCCAGCAGTTTCGCTCAGCGGAGTAAGACGCTTATCCAGGAGGAGGCGCGGTGGAACCGTTTTCATGAGGAT
TGCCGCTGAAAATGATGCGGAATATGTTGAGCCAGCGGTTGCCGCTTTCGCGACGAAGCGGCTCTTATACAAATGG
GCTTCAGATGCCCGAGGCGAGTTCAGCAGTGTGCAATGGGAAAGCGGACAGCTCATGGTGCTGTGCTGCTCTTCC
GACTGAACCAGGTTTCAAGTGTGTTGACGTTCTGCTCAGCGTTTCACTCGGTTGTTGGGATGGGCGAGTGGGCGCA

FIGURE 2N

GTGTGAGCTGACGCTAGCGAGGTAAAGAGTGGATTGGGACTTTTGACGGCCCGCTGGCTCTGGCTCGATTGGCCAGACG
GATAGAGCCCGTCTTTCAACGAGCTCTGGATGAGAGACAAACAAATGGCGAGGCTGATGAATTCGGACTGTCTTGTTCG
ATTCCGGCGGTGATTTCCTCTGTGAGGATCATCGCGATGTCCAGCAACTCCATGCTGTCTTCAAGCGGACCTGATTTC
CTGCATCAGGTGCTCCAGCGGAACCATATCTGGCGCTCGATCTGTTTCATCGAGCTGCCGCGAGCGACCGCGCGGGT
GGATCTCCAGCATCAAGCGCTTGGTGGCGGTGTCTCCGGCGACAGACCGAAAAAGCTCTGGATCATCGTTGTGGTGGCT
CCCAGCTTCAGTGCCTGCTCGATCTGCAGGTCTTCTGCGCGCTGCGCTCGGCTGTGGCCAGGATTTTCTCCATGACGTC
GATGTTCACTCTGACATCTACCCAGGAACCGGGGTATTGGTCAGGATGCTCATGACGGCAGGCTGTGGAGCTGAGCCA
GAATCTCCGGCCGAGCCCCATCTCGGCACAACGCTGAGCTGGCCGTTGCGCAGGTGGTTGAGCACCTGCGTCAGGACC
GCTTCATTGATCGTCACTTCGACATGTGATCCCTCCCAAAATCTGTCGAAATCACTGGTCAGTAGCTGCTTGCTCGGA
GGCGGCGCGGCGGAGCTCAAGGTCAACAGGCGTGGGCAAGACGAATGATCCGGAACAGTTTCAACAGGGCTTCGTCAT
TGAGACGGCTGCTGAGCAGGATGGCTGTCACTTGTGGTGGTTGGCCAGCAACTGAGCGAGGCTGAGTGAAC
TGAATTCACCCAGTGCATCGCGCGAGGTCTGTTGATCGTGAGTGTTCAGAACGACTTGACCTGCCGCTGACCGCTTG
AAGCAGCGTCATGAGGTGACGCGCTGAATCAGTAATCTCATGTCTTCAATTGGGCTCGCGATACGTGTAACCCACACCGC
CCTGAATCTGACGATCTCGCTGGGGCGCGACATGACAGCGATCTTTCAGCCAGATCAGCGATGTGCTGTGCGCAGG
ATCTCGGGGCTGCTGAGCAGCTCGATCTACAGAGATCGGTGATCGGGTGCAGGCCACCGACTTGCATGGGATTGCG
CTGCAGAGCTGTTTCATCGAAGACAGGTAGATGTTCCCTTCAGGCGCGGCGAGCGGTTGGCGAGTCTGCTGGATCTTAG
TCGATACCGGGCTCAGCATATGGCTGCCAAGACCGCGCACGCTCTCATCGGCAATTGCTTTTGTGGACAGGTGGT
GGCGGAGCGCGGGGGAGGGGATTACCGGGTGTCTTGTCTGACGTTTGGGATCTGCGCGGGGCTCTGTGT
TTGTGGGGTCTGCTGAGCAGCAGCAGCAGGTTTTCTGCGCAATTGGTGGGAGTTGAGGTGGTGGCGAGGTGGGCA
GGTCGACAGAGTCCGCGCGCTGTCTCTGCTGCCGTTGGTAATCTCGAGCAGGATTGGTCATAACGAGGCCAGG
GGGCGCTTCATTTGGTCGATGAGTTCGTCTGGAACGCTCGAAAATGAATTCGTGAGGCTACTGTGGAAGATTGCCAG
GGTGTCTTGAACAGCATTCGAAGTCAACCCCTTACCAGCGTATAGACGTGCCAGCAGGAGGATGCCAGCTTCTGA
GTGACGGGTTCTCGATTTGCGGTTTACGAGCCCTGAATACAGCATCACAGGAATCGCCGCGCAATACCTCAATA
GTGTCTAACATCTTACTGATATGGGATTGAGATACAGGCTAGCCGCTCGCTTTTAGCCGACGTGCCAACTCGCGCTGGGA
AATGCTTTCACCGCGCTTCTTGTTCGTAAGAAATTTGCGCTTCTGACCCCCAACCGCGCTCGATGAACCTGAGGT
CGCCCTTCAGATCGTTCTCGGCCAATGACCGGTGAGCGCATGATTCGCGCGCTGCTTGTCTCCAGGCTTGAACAGG
CAGTCGAAGCTGAATAGCGCTCGCTCTCGGCTCTCTTGTAGAGGTGTTGAGAATTTCCAGGCGAGTGTTCGCGGCT
GCGAATGCGGTATTTGTCTCTCCAGGCGCTGAGTCACTGGTGGCGGCTATCGAGGCTCGATGTGCGATCGATTCTT
TCAGCTCATCTGATTTTCGGGTTTCGAGTGGTTCGGGGGTTATCGTGCCAGGGGAGGACGTCTGAGGATGTGACGCTCATG
GGCTGTGCTGCTGATCGGCTCGGCAACTGCTCCAGAACCTTGGGCGGTGCGGTAACACAGGGGCAAGTAGCCGCTCTT
CAGATCTTTCGGGTTGATCTTGGCCATACTCATGCTCCGCCACTGAGGCGGTTCGCGCGCGCGAGCTCAGCTGACG
GCGCTGGTTCTTGGCGAAGAGTTTCGTGACCAATTGCTTTCGCGCGAGTCTTCTTTCGAGCGGTTGCTGCGCTCCA
GTGCGTAGACGATCTCGCTGTTGGCCGAACGGTGGTTATCCGCGGCTTTTGTCTTAGCTGATCGCGCAAGCCCAAGGGC
ATACGGAACAAGCACTTATCTTCGGTTTCTCGGCCATGCTCATTCTCAATTTGGGCTTTCGCGGCTTTCGAGTGGTCTG
CGGAGTTGAGTGGGAGGCTGCTGATCTGCAGCGTGTGCGGCTCTCTGTCGATCTGATCTGAGCGCCAGCAGGTGCGC
CTCGAAGCTGATCGACAGGCTTTCGCGCGCGCGGTGAAGCGCGGAACGTTGAGGTTGCGCTTGTCCGCGGGGATT
CCGCGCAGAGGCGGTAGTCTTGTACCGATGTAGTCGTAAGACCGCCGCGGTGCTGGTCTGCTCATGATTCGGAAGC
GCGTCGAGGCTCATCGGCTCGCGGATGCGCGCTGCGAGGTGCGGTAGTCGACACGCTCTCGGTCTTCTCGCGGCGCTG
TTCTCTCGCCATGCTCTCGCTTTCACGAATCGCTGAAGGCTTTCAGCAGGTTGCGGCTCTGCTCGCGAATCCACCC
CTTCTCGGAGCGATGAAGTCGCGGAATAGTCGAGACCTTCTCTCCGCGCTTTCGCTTTCGATGAACGAGATGTACTGC
TTCGACTGCTTGTGTTGCGCCATTTCGAAATGTTGATCCGCGCGGCGAGGTGCAACTGGCCGAGGTCCAGGTGGCGCA
CGGGTGACCTTCAGCGACTCGTTCAACCGCACTGCTGCTGTTGAGCAGCAGGCGATCGCCAGGTAGTCGCTCATGC
CTTGTCTGAGTGGGCGAAGCAGGACGTGGCGCGGTTGAAGAGATTGAGTCTCTCATCAGCTTTTGCAGGTGCTCGACC
GCTTCGCGGCTGAAGCCGACGAAGTCGCGGTGCGCTCCAGTACTCGCCAGCGAGCGCTGAACGGTAGGCCCCGGA
CTCGCCCTGGAAGAAGCCCCAGGCTTGTCTTGGCGTTGAGCTTTCGTTGAGGTGCGCCAGCAGGTTCTCGATGGCT
GGGAGTCCGCGAGCTCGGCGTGCAGCGCTGAGCAGCGCGGGTGGCGTTCGCGCTTCTTCGATCAGGTGGACGATG
GCGTGTGATAGGCATCAGAAAGACTCTGGTTGAGGAACGCTGCTGTGCGCTCGACGAGCAGGGCGCGGCTATTG
AGCTCGCGCGTGCCTGCTTGGCGTGAAGAATGAAGTCAAGGCGCAGCGGATCTTGTGAGCTTCAAGTGGCGCC
TGCGGAGCGGTGACGATACGGCTCAGGCTTCTGCTTGAAGGAGCCAGCCGCGCAACGCTTTCGCGGCGCGCGAAGGA
AACTACGCACTTCTGCTCTCTGCTGTTGAGTGCAGCCAGGTCTCATCTCTGCAAGGAGCGGGGAAGGCTATCGGT
GGCGCGCGCATCCAGGAGACGTTTACAGCGCTGAAGGCTGCGACGATCGCTGCGCGCTGACGCGAGGTGCTCAT
GCGCGAAGCAGCATGCGGACCTCTGTGCGCAGTTGAGTGCATACGTGCTGATGCTACTGGTGGCTTGAACGACCA
GGTGACCGGACCGTGTCTGAGGCGGTTTCCCACTCTTCTGCTGCGGAGAAGCGATGGCCACGACTACACTCCA
ACCGCTGGTCCAGTAGCGGACTTTCGCGAGTCAGGCGAGGCTGTGCGCGCTGCCATCCAGGTGATAGCCGAGTTCA
GACAGATCCACGAAGGACCTTACGCAACTTGAGCTCAAGCATCAGCTCGACTTCTGGAAGTGTCTGAGCGCGCGTAG
GCTGTGAGCTCAGGCATCGCGCGTACCTCTTGGGATTTCTACAGGACAGCGCCAGCGGCTCTGATGCGCTTTTGT
GGTGCCTGCTGTACCGGATACGGGTGATGATGCTTGTCTCGCCAAACAGTCTGTCGCAAGGCGGGAGGCTGCC
ACCTGTGCTCCAGGTGACGCTGCGACGCTGCTTCTGAATGGTATTGGTCCAGAAACCGCGCGCATCGAATTCAGCAT
GCACTGATAGGTCTGGGACGCTTTCATGACTCGGCTCTTCTGCTGAGTGCATCAGGACTGAAACGCGCACAGTTGCGG
CAACTCGGCTCATGCCGAGTATGTTGGCCAACTGATGACGAACTCGCGCGCGGCTCTTGTGCTGCTGCAATG
GGAACACGCGGTGAGTGTGGGAGTCTGAGTGCAGCAGAGGTGCGGACAGGCTCACTTTCAGCGAGCAGCTG
AACAGGTGTGCGGGATCCATACGACAGGAAGTAAGCCGAAAGCAATTGCCCCGAGCGCGCTAGTGACACAGAGT
CCCAGCTGGTGGTTCGAGATTTGATGTATCCCAAGTTCCTCCAGTCAAGAGCTCTGCTCATCGAATCGGGTGGCA
CCGTGCTCTGCTCGCCGGAAGGCGGAGATCACTTGGCTGAGGCGCTCAGCTCGATGGGTTGACCCCAAGCTGATTG
GCATGAACCACTGCGGAGTGGTGAACGCGGCGAGGCTGATACCTCGGAAAGCGGGCTTCACTGTTTCATCAAGACAG
CCTCAACTGATGGAAGTCTGAGTTCGCGAGCAGGGCGAGCAACACCTCGGCTTCTCTTCTGGGGTGTCTCGGCG
GGGAGAACTGGCGACCAATGGTGTGGTGGTGTGGTGCACAAATTCATGCACTGGTGGGTGACAGGGGCTGGGAAGT
GTTTCATGATGGGCTCTTTCAGGTCCCGGCGCAGGTGCGCAAGCGGGTCTGGTGGCGATGAAGCGGTGTGGAC
GGTGGCGATCGGACGTTGCGATCTTGGCGATGATTTCGCGATGATTTTCGCGATGCCCTTGGCTCGGTGTGGGGTGGTGA
CGTGGCGGTACGGAACATGATCAGTCCGCGCTCTGCTGATGCGCGCTCTCTGATGGTCCGCGAGTTCGGTGGC
TTGTTCTGCTCGGTTCTCTAGGCGCGATTAGCTGGGATAGAGCTACGACGGGACAGTCCATCTCTTGGCCAGCGCTT

FIGURE 20

CAGCGAGCGGGAGATTTCAGAGATCTCGGTGGCTCGATTCTCCAGGCCTGGGCAACTCATCAGTTGCAGGTAGTCGACCA
ATATCAGCGCGGGGTGTCCTGACTTCTTGGCCGCGCGGCAACCTTGGCGCGCAGCTCTGTGCGCGTGAGGTTGCCCTGA
TCGTTGATGACCAAGCGGCTGCCATAGTCGTTTATGCGCTGGATCGCGCCAGACAGTCTGGGCCAATCCTCTTTCGACG
TTGGCCCTTCATCAGCTTGCCAGGTCCAGGTGGCCGAACAGGGCGGCAAGCCTGAACAGCAACTGCTCTGCCGCATCT
CCATGCTGTACACTGAACAGACTTCTGTTGGTTCGCTCTGGAGCGCGGTGTCAGCAGGTTGAGGGCAACACAGCTTTTG
CCCATCGCGGGGCGGGACCGACATGATGAGATCCGACTTCTGTAGTCCGCGGTGAGTGCCTGAGATCCTTCAGGGC
AGTGGGACCCCGTTACCGTCACGTTGTTGTTGAAGCGGTAATCGATGGTGTGACGATCTTCGTGAGACTCTTGTTGA
TATCGACGAAATCGCGGTTGTGGTGGTCTGGCCAAGGGCCGAACAGCTTCTGCTCAATCTCTCTGAACCTCAGAGGG
ATTTGCTGGTGGTTCGAGGCGGTCTGTTGACGTGGTGCCTGAGAGACATCAGCTGCCGAGGTGTGCCGATCGCGAA
CGATCTGCGCGTAAGCCTCGATATTGGCCAGGAGGGCGTGTGTCGCGCAGCTGGCCGAGGTAGGCCAGCCGCCAGCT
TCTGGAAGGTCTTCGATCGCTTCCGACACAGTCAAGCATCAACAGGAGCATCTTCGCGGCCAATCCTGATGGCGGT
GAAGATCAGCCGATGCTCATGCCGAAGAGTCTCTCTGCAACTGATCGCCGACATGTCCCATGCCGCGTGTGTCA
GCATCAGGCCGCGAGTACCCCTTCTCGCCCTCGACGGAGTGGCGCGCGTAAGTTCAGAACACTCACTTCAGAGTCTCT
CCCGCTGATCTGACAGAACTGCTTCTCTGCGCCGCGGGAGCGGATGATCTCAGCGCTTCACTTCCACTTGGCG
GACTCGATGATCACCAGAACGACATGAGCCACCGCTTTCGCGCGCTCGATATCCAGGGGCTTGTCTTCGTGAGACTT
TTCGATGGTGGCAAGAGGTGATTCCGACAGATCTCAATCTTGTCTTCATGCTCGACCTCTGAATGGTGGCTTGTAGC
TTGCTGATGACGCGGATTGCGCTCTTCACTCAGGTGGATAGCGATGGATGGAGTTGCGGCGCATGCGTTGCGCCCGGT
GATGAGCTCCAGGTTGTGATGACGAGCTTCTGCTTGTGTTGTTGCTTGAAGCAGACGCAATGGCCGGTGGGATGGGG
CGAAGTGTCTTCCAGAGCAGGATGTGGATGCTTCACTCCGCGGGGATAGCCGATTCAGAGATCTTCGTTGTC
AGGTAGCCATCAGCGCTGACCCGCGTGTGCGCACTGGGAGCCATGTGTGGGCTTTTGCCCTTCTTGAACCTGCGTTT
TGGTGGCGCTCTCGCGCGGGCAGCCCTTCACTGCGCAGTTCCATGGGCTGTTTCCCTTCTTGAACCGAGTTGCGCTGC
CCAGTCTCTGCCGATCTTCTTGGAGAACTCAGGGCTTTTTTTCAGCCGAGCCGATACGCTCTGGAGCAGATCTGCTGG
AGCGTTTTGTTCAACCTGGCTGCCAGGACCTCGTTCCGATATCCGATACAGGCGCAGCAGCAGCTTCTGTTCCGCG
GGTCCAGGTTTCCGCGATGACGCGCTCTCTCGGTTGCTGAATTTGCGCGGGCAGCAGCAGGCGCTCTGAGGCGATGG
TCGAGAGAGCAGAGCGAGTCAATGGGAACCTCCTGCTCGAGCGCGCTTGAATCAGCGGCTGACACCGCGGTGACGCG
GAGTGTTTTTGCGCTGCTCCAGCCTGCTTGCAGCGAGGCTGGTATAGCGCAAGGGCTTCGTAAGCGCGCGCGCTCC
AGGTACTTCAACGCGGATGGTGTAGTTCAGGTAGGCTTGTGGTCTGCTTGAAGTGCAGTCTGCTCGCGCTGCGAGAAA
TTCAGTACAGCGAAGCAGTGCAGAGAGCCAGCTTCGACGAAGAGCTTGTGCGACGAGCTTGGTTCGACAGCTTGC
AGCGCTTCTGTTGGGCGACATGAGCCGAGCGCTGCCACAGTTGGTGGTGGAGCGTAGAGTAGGCGTGAGCGGCCAGC
TCAGGGCGGATCCCTCGGCCATGAACCTTGAACGCTCAGCCGCTGGCATCGGTATGCCGCGAGGTAGTCGACGCCGAA
GGAAGTGGCGCAGATCCGCGCAGTTTATGCGAGCCTAGGGGAGGAGCGCTGGTTCAGTGGGAAACGACGCGGTTT
CCACCGTGTGGGCGAGATCTCCGATCGCTGAGGTGCAACTGCCGCTTCAAGATCGCGGCTGGCGCAATGCGGTTTCC
GCTCGTTGCGGCTGGCCCTTGGCATTGCCAGGGCCATCAGCTTCTGACCTTGGCGAGGACTTTTTGCTTGTCCAT
TAGCGGCCCCCTTGCAGATGGTGTAGCCCTGCTCGGTGGCGGAGCGGTCCAGCGCTCGATCTCGGCCAGCACTAGGGC
GCCGCGCGTACAGGTTTCTCGCGCGCGCACTTAGGCTTCCACACTCTTCTGCTCCAGGGCCAGAACAGTGCAGAGCCAT
TCCGAATGCCGCCCGCCAGTGGTGGCGATGCCGCGCGATGGAGCGGTAGCAGCGCGCCAGGCGGATCAGGCCA
CGCTAGCGTCTGCTGCTGCTGATCGAAGCTTGCAGCGTATCTGGCGCGCGCTCGGCTGTACATCGAGCCAGGC
CTGCGGAACAGGGGAGAGGTTGGAAGTTGGGTAAATTCGCTTGTGGCTGTCTTACCGGATCGACAATAGCCAGGGCTG
CTACGTCGCGCAGTGCAGCTCCAGGCGCGGAGCAGGTTGGCCATCAAGAACTGCTCAGCTCGTCCGCGTGGTTTTTT
GCGCTTGGGCGTCTATTCCAGTGCATCTGTGCTGAATACCTCGCTGTTACCGGCTGGCTCGCAGTAGGACGCGCCGAT
GTGAACCTGGCAGTCCAGGCATTCAATGGCAGTTCAGGCATCGGTGTCAGCAGCTCGGTTCCAGCATTGCAATGT
TCGTGCTGCCGAGAGCGGGCAGGGCTTGAAGTTGGGTCATCATGGCTGCTCTCTGCGCGTCTCGGTGCTGGGCT
TAGCTTGGGAGGAGCGCGCAGCAGCTCGGCCAGGTCCGATAGTTTCGATGGTCCAGAGCGCAACTGATCGCGC
TCGACAGGTGGGCTGAATCATAGGTTTATGCCAGGCTCCACGCTCATGCCGAGCACCGCATTTGCTCGCGCGG
AGTTGCTGGCTTCAAGTGTGATGAGCTGCCGACGAAGAGCATCAGCGCTTCGATGACCTCAGAGTGTCTGCGAGG
TTGGGCGAATCGGCCAGGTGATCGATGGCAGTAGCTATGTTCCAGATCGCCCTGTTTATGTTCCGCTGGATCACTTGA
TCGAAGTGGCATTGGCCGCGATGGCCAGTTGCGATTTCGCTGCAACAGTAGGTCGTGGGCGATTGGAGCAGTGGCTC
TGAGGGAGTGTCCACAGGCAAGCTGCTGAGCGGAACCGAGCGAAGGCGCTGCAACATCGCAGCACCTTCAGCCTC
GATTTGCAGATTCTCGAAGCACCGCGCGAGCGCTGCTTGGCCATGATGCTTGCAGTAGGCGTCAAGTGGTCT
CGCGCATCAAGCTCAGTAGGAGGTTGTTGCGGTTCTGAGGTTTATATGAGGCTTCCGAGCGGATTCAGAGCGTCTG
TCCAGAGCCGAAGGGTTTCTGCTCTCGCTGGTCAAGCGATCTGTGCGGTGACGTGAGCGACCTTCTGCTTGTAGAGC
CGCAGGCGCTCTCGACTTCACTGCCATCGTGTACGCGATTTCGTTAGTCAAGTGGCTAATGAAGTCAATCGCAGGCT
CCCCACCTGACTAGCGAAATACTCGGCTTCACTGCCCATCCGATGAGCTGCTTGGAGCTCCAGCCAGCGTTTGAAGT
TTGAGTGGAGCTGGTGGCTGAGTACAGGTGATCGAACACCTTGGCCCGGTTGATGATCAACGCGATCTGCAGGATTCC
AGCCGATCTTCAGCGGATTTCCCGGAAGCGGGTAGTTTCAGGCGAGCAGGTTGCTTGGCGAGGCGCTCAGTAGCTCGCG
CCGCTTCTTGGCTTTTTTCTCGTAGCGATCAGCTTTTCTTTCGATGCTCGATCTGCTTGTTCACCGAGTCGAGCAAGC
GCGTGTGCTTGTAGGAGCTGCTTTTGTTCGAAGCTCAAGTAGCAACCGAGCTTCCGAAACAGGTTCGATACCGCCACCA
GTCAGCTTGCATTTCTCGCGTAGCGCGGTCATTTGGCCTGCTGGCTCCGAAGGTGCGCGCGCTGCCGAACCTGTT
CAGTTCCAGCACAAAGTTCGCGCTCGAGCTCGCGGCTTGGGTGATAGGAGTCTTCGCGATGTGAGCGCTTCCCTTGA
AGTGTCTCAAGGCGCTTCCGCGGTGAGCGCAGGAAGCGGTGAGTTCAGTCCGCGGAAAGACCTCGATGGCCAGGTTTCGAA
TGATTTCCAGCGCGGGCGCTGATGTGCGATTGGAGGGTTGCCGCTTTCAGAGCGGTGCGCTGGTAGCCCGCGATGTCT
GCGTTGCGAAACAGCAGCGCATCCGGGACAGTCTGTTTCAGAGCAGAAATGTTCTGATGCTCATGAAGATGGCAGCGAC
ATTCTCGTGAATTTGCCGGGAGTCATTGGTCTGGTCAGGCAAGTGTGATGAGATCTGAACATTGGGGATCCGATGCCGA
AGCGCTCAAGGGGCTAGGCGGTGAGCATTTGATGGTGGCGGTTGAACCTCAGGCGGTAAGCATGTTGGGTTGG
AGGGGGAAACAACAGGTCCGATGCAAGCACAACCTTTCAGGCAAGCTGAGCGCGCACCTGGGTGTCGATCAGCAC
CAGGTGATGCTTCTTTCAGAGCGGCGATCAGGTTCCGAGGCGTAGCGCGCATCGGGCGCTGGAGCAGTAGGTTGT
TGAGCTGATTGTTGCTGCTGTTGAATCAGACGTCAGATTGGGGATAATCGTCTGGAGATGATCTCGCCGGGTTCC
GTTATGTTGGCGGCGAGGCTGTAATGCGGCTTGGGCAACTTCCGCGAGCTCGTAGTACGAGGATAGGAGGGCTG
GACGGGTCAGATCGATGAGGAGGTTTGTATGCTGCTGATCGCGCAAAATGCACCGAGGTGGCGGCGGTGGTGGATT
TTCCTACACCACTTGTGGATACAAACGAAGTCGCTTTCGATTCATGTCTACAGCTTCAATGTGGGGTCAATAGAG
GCTGCTAAACAGACTCGTTGTAACGGGAGTGAAACCTTCTAGAGAACTACGACAGTGTGCTTTTAAACCAATTGG

FIGURE 2P

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 18 of 118 Customer No.: 21559

TCGTAGGTTTTCGAATCCTACACGACCCACCA

FIGURE 3

RL024

DNA sequence (SEQ ID NO: 3)

```
GTGGCGCTGACCGGTAATCCCCTCCTGAAATTGCTGGTCGTCCCGTCGTGATCGGCGCCATCCTGATCGGCGTGAGCAT
GATGGGCAAGAAAGAAAGTGCAGTCACAAGGCGCCGCAACCCGACGGTAACGTGGAAGAAGCGGCAACCCCTGGGCA
TCGACGGCGACACGCCCGCGACACTACGCACCATCGTGGCGGAAAGCCGGCAGCTCAAGGACCAGATCAGCAAGGTG
ATCCAGGAGAATGACTCGCTCAAAGCCGCAATGAGAACCTGCAGGGCCGCTGCGCAACATCGATCAGAACATCGAGCA
GAAGCTCAACAACACCGCCAGGAATGCAGCAACAGCAGGAAAACCGTAGCCAGACGATCCTGGACCAGGTACAGAAAC
GGCTCGAGAACCTAACCCACATTCCCGAGGCCGTGACACCGACCTGCCGTAGGATTGCGCGTGCGACCAAGGATGGC
CAGCACTTTCAGGGAGCGGGCTCGTCTTCATCGGATATCGTCTGGATCGAGCCCCAGGACGCCCGCGGGTTGATGCCAA
TGCCACGCCGTGGCGCGCGGCTCCACCACCCAACCGAGCGGATTGAGCTTCCCGACCTCCTTCGGCAATGCGGTGATC
GCGGACAGAACCGCGTGGAGCGGATCGATGACGGGCTGCACCCCGTCGGCCAACAGCGATCTGACCTGGAAAACCGCAAG
CTCCTCCGCTAAGACCTACACGCTGCCGAGAACTCGACGCTCATGGGCTCGGTGGCCATGTTTGCCTGATCGGTCTGT
GCCGGTCGACGGGACGGTCAATGATCCTTACCGGTTCAAATCCTCATCGGCCCGGACAACCTCACCGCCAACGGCATCG
AGCTGCCGGACGTGCGCGGCGCGGTAGCCAGCGGGACCGCTCGGGCGACTGGACACTCTCCTGCGTGCGTGGGCAGATC
CGCAGCCTCACGTTCTGTTTCAACGACGGGACCGTGCGACCTTCCCGCGCGCGGCGGAGGAGGTGAATGACAACAGAG
CAACAACAACAGACCGCCAGCGCGGACCGAGAAAACCATCCAGGGCGGCTCGGCTGGATCAGCGACCCCTACGGCATCC
CATGCATCGCCGGTATCGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGCGCGG
ATTGCCAAGCTCCTGGACGCCGACGAGAACCAACACAGTACCGTCTTCAGCGGCAACGGCACAGCTTCGGGACGACCGG
AACCAACAGCAACTCGGCCCTCAACAGCATCCTCTCCGGCGGCGTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACG
GGGAGGCCTTCGCCCGCGTCTACGTGCGCCAGGTGCGCGGGTGCAGTGCATCTCGATCAGCAACTGGCGATCGACTAT
GAACTCAAGGGCCGCAAGGTGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAA
```

Protein sequence (SEQ ID NO: 127)

```
VALTGNPLLKLLVVPVIGAILIGVSMGKKESAQSQAATPTVTSEEAATLGIDGTPADTLRTIVAESRQLKDQISKV
IQENDSLKAANENLQRLRNIDQNI EQKLNNTAQELQQQENRSQTI LDQVQKRLNLTHIPEAGDTLDPVGFVVRPKDG
QHFQAGSSSSDIVWIEPQDARAVDANGQPLAAGSTTPSGFSFPTSFNGNAVDRGQNALERIDDGLHPVGQQRSDLENRK
LVRKTYTL PQNSTLMGSMVAMFALIGRVPVDGTVNDPYPFKILIGPDNL TANGIELPDVAGAVASGTASGDWTLSCVRGQI
RSLTFVFN DGTVRTFPAPAEVNDNQSNNNQTASADQKTIQGGLGWI SDPYGIPCIAGDRRSNAKEYLGNQSLLLTAAGAG
IAKLLDADENN TSTVFSNGTSFGTTGTNSNSALNSILSGGVS DIRQWMNKLYGEAFAAVYVQPGARVAVHLDQQLAIDY
ELKGRKVDYSSGAHAHADLD.
```


FIGURE 4

RL025

DNA sequence (SEQ ID NO: 4)

ATGATCCGGAAGTCGACAGGCTCGCTCTTGCTAATGCTTGCCCTACCCCACTGGCCCACGCGGTGGAGATTCTGCGCTG
GGAGCGCATTCCGTTGGCCATTCCATTGACGGTCGGCCAGGAACGCATTGTTTTCGTCGACAGAAACGTGCGAGTTGGGG
TTCTCTCGGGATCTGCAGGGCAAGCTGCGCGTCCAGAGTACCGGCGGCGCACTCTACCTGCTCGCCAACGAGCCGATTCTCT
CCAGCGCGCTGCGCTACAGGACGCGACCAATGGCGAGCAGATGCTCATCGATATCGCCGCCACCGAAGCAACGGCCGA
CCAAACAGCCGCGCGAGCCGGTCAGGATCGTCGCGGCGAGCCAGTGGATCCGCATTATGGCCAGTCCCGGGAAGCCAGC
CATCGGCAGCAGCGAAACAGACCGAGCACGCAGAAGCACCGAAGGCCGTGCGCGCGAAACGCCCGTCCCCGTGGTTCTG
ACGCGCTATGCGCGCGAGATGCTCTATGCCCGCTTCGCACGGTGGAACCGGTGGATGGCGTCGGTCAGGTGCGCGTCAA
GCGACAGCTCGACCTGACCACCCTGCTCCCCAGCCTACCCATCACGGCTACCGCCTTGGGCGCCTGGCGGCTGGACGACT
ACTACATCACGGCGGTGAAGCTGCAGAACGCCAGCGCCAGCACCTGGCCCTGGATCCCAGGGACCTGATGGGCAATTTC
GTCGCGCGGACCTTCCAGCACCCGTACTTGGGGCCCCGGGGCGACGCCTCCGACACCACTACCGTGTATCTGGTGACGCG
CGGCCGCGGCCCTTGCCGACGCGCTCCTGCCCTCCTCCATCAGCCAGATCGATCCCAAAGGAGGCCGTCGTGGCGCTGACC
GGTAA

Protein sequence (SEQ ID NO: 128)

MIRKSTGSLLLMLALPTLAHAVEILRWERIPLAIPITVQGERIVFVDRNVRVGVPRDLQGKLRVQSTGGALYLLANEPIP
PARLRLQDATNGEQMLIDIAATEATADQQPREPVRIVAGEPVDPHYGQSREAQPSAAAKQTEHAEAPKAVPRETPVPVVL
TRYAAQMLYAPLRTVEPVDGVGVQVRVQRQLDLTLLPSLPITATAGAWRLDDYYITAVKLQNASAQHLALDPRDLMGNF
VAATFQHPYLGPRGDASDTTIVYLVTRGRGLADALLPSSISQIDPKGGRGADR.

FIGURE 5

RL026 : DNA sequence (SEQ ID NO:5)

ATGAGTTTCAGAAAACACACTGCGCAACAGCAGGCACACATCAACACCTTCCGGTTCATCACCGGCTTCCTGTGCATGGT
CATCGTTGTGCTGGCCTACTGCGTCTGGGAAGCCCGTAAGGACCTCTGGATCCACATTCCGCCCGACTTGCGCTCAGGAA
GCACCCGGTTGTGGTGGGACATTCCGCCAGAGAGCGTCTATGCGTTCGGCCTCTACATCTTCCAGCAGGTGCAGCGTTGG
CCCAAGGACGGCGAGGTGGACTACAAGGGAACCTGTTCCGCTACGCTGCCTACCTCACTCCCTCCTGCAAAGTCTTCCT
GGAGAAAGACTTTGAGTTTCGTGTAACGCCGGCGAGCTCAGGGGTCGCGAGCGCACCACTCGGAAATCCCCGGTCGAG
GCATTGGCGAGAGCAATGGCCGCGTGATCCAGCACTCGATCAATGACTGGACCGTCAACTTGGACATGGACAGCACGGAG
TATTACGCCGGCGAGAAGATCAAGCGGGCGCTGGCCCGCTACCCGTTGCACGTTATCCGCGCCGACGTCCGACCCGAAAC
CAATCCCTTCGGCCTGCAGTGGGACTGCTACTCCGACACGCCTCAACGTATCGAGCTTGAGGAGCCGGCCGCCCCACCA
AGCGGGAGGGAGGTCTATGA

Protein sequence (SEQ ID NO: 129)

MSFRKHTAQQQAHINTFRFITGFLCMVIVLAYCVWEARKDLWIHIPDLRSGSTRLWWDIPPESVYAFGLYIFQQVQRW
PKDGEVDYKGNLFRYAAYLTSPCKVFLEKDFEFRNAGELRGRETTSEIPGRGIGESNGRVIQHSINDWTVNLDMDSTE
YYAGEKIKRALARYPLHVIRADVDPETNPFGLQWDCYSDFQRIELEPAAPTKREGGL.

FIGURE 6

RL027

DNA sequence: (SEQ ID NO: 6)

ATGCCCCGAAGAACATCTGTTTCAGGATGGAACCCTCAGCTTCCTGCCGACCC
GTTTGAACCGGCAACCGGTAGTCATCGG
CGGCCTGACCGCAGACGAAATGTGGATCACGGTCTTCACCAGCGGAGCAGCC
GGGTTTCGTTCTTGGCATCCCGGCTGCCT
TGGTCGCAGGTAACGCTGCCTGCATTCCACTGGGCGCGCTGCTGGTCGGCGC
CCTCGGCCTAGGTATCGGCAGCCGCGTC
CTGCGGCGGATGAAGCGGGGGCGGCCCGATACCTGGTTCTACCGCCAGGTGG
AGATGGCCCTCTCGCTGCGCTTTCCCGT
CTTCGGCAACCGTCGCCTGGTTACGCGCTCCGGCGCCTGGACCAGTCGACGC
ACGGAGTCCCCATGA

Protein sequence: (SEQ ID NO: 130)

MPEEHLFQDGTLSFLPTRLNRQPVVIGGLTADEMWITVFTSGAAGFVLGIPAALV
AGNAACIPLGALLVGALGLGIGSRV
LRRMKRGRPDTWFYRQVEMALSLRFPVFGNRRLVTRSGAWTSRRTESP.

FIGURE 7

RL028

DNA sequence: (SEQ ID NO: 7)

ATGCTGAAACTCACCTCCAGAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCATCACTTTGGCGCTCCCCGGTATCGC
CTTGGCTGCACTCCCCAAACCGAGGCACCTAGCCGTGGGGAGGGATCGGGCATCATGCAAACCATCCAGAACTTCGGCT
ATGACGGAGCGATGCTCCTCGCGCTGCTCATCTGCGCGGCTGTCTTTCTGGGGGTCGCTTGGCATACCTACGGCACCTAT
CACGCCATCCATGACGGGAAGAAGAAGTGGTCGGATCTCGGAGCGGGCGTAGCCGTAGGTGTCGGCCTGCTGATCTTGAT
CATTATCTCGTCACCAAAGCCACCGCCATCATGTAA

Protein sequence: (SEQ ID NO: 131)

MLKLTLQKLSALCQSLAAITLALPGIALAALPKPEAPSRGEGSGIMQTIQNFGYDGAMLLALLICAAVFLGVAWHTYGT
YHAIHDGKKKWSDLGAGVAVGVGLLILIIYLVTKATAIM

FIGURE 8

RL029

DNA sequence: (SEQ ID NO:8)

```
ATGAGCATGAGCGGAGCCCAGACATCAGCGTTCCAGGCCGCGCTGGCTTTCCCCCATCGGCCGGCGAGGGACTGTTTCAT
TGGAGCAGCGATGACCTTCCTTCTGCTGTGGTCCGCTGGGCGATGTACAGCACCTGGCGCGGCTGGGCCACCAACAACC
TTCGACAGCGCCACCGGTGGCGCTTCCGCGATCCCGGATCTTGGTCCTCCTCGGCATCACCTCTTCTTCTCCTCCTCAGCT
GACCCATACGGAGACACTCATGCTGAAACTCACCTCCAGAACTGTCCGCCCTCTGCCAGAGCCTGGCCGCCATCACTT
TGGCGTCCCCGGTATCGCCTTGGCTGCACTCCCCAAACCCGAGGCACCTAG
```

Protein sequence: (SEQ ID NO: 132)

```
MSMSGAQTSAFQAAAGFPPSAGEGLFIGAAMTFLLLWSAWAMYSTWRGWATNNLRQRHRWRFRDPGSWSSSASPLSSSSA
DPYGDTHAETHPPETVRPLPEPGRHHFGAPRYRLGCTPQTRGT.
```


FIGURE 9

RL030

DNA sequence: (SEQ ID NO: 9)

```
CTGATCTGCACGAGATTGCGCGTGAACACTCCACATCCATCCCTTCGCCGAAGCTGCCTGGCCGTCTTGGCCTGCAGTGC
GCTGGTCGCACAGGGAGCTTTCGCAGCGAGCGCCTCCGAGCAGGCGAACCTGGAGGTGATGATCCGGCAGCTCAACGCCC
TCGAGGACACCGCCCGCCGAGTGCCCGAGGGCGCCGATGAGCCCGGACAGCGCTTCTACTTCGACTACCCGCGCCTGGCC
GCTGACCTGCAGCGCATCCGCCAAGGCCTGCAGGACTACATGACGCCCGCGCCCAACCGCGTGACCCCTCCGACTT
ATCAGGGAATTACACCCTGCGCGGAGGGCCGATGCCATGA
```

Protein sequence (SEQ ID NO: 133)

```
LICTRFAVNTPHPSLRRSCLAVLACSAALVAQCAFASASEQANLEVMIRQLNALEDTARRSAQGADEPGQRFYFDYPRLA
ADLQRIRQGLQDYMTSPRAQPRDPSDLSGNYTLRGGPMP
```

FIGURE 10

RL031

DNA sequence: (SEQ ID NO: 10)

ATGAGCATAAAACAGCCCTTCGAATACCATGTCGAGAACATCGTCATTCCCTACAAAACCCTCACCAAGGGCGTCGCGAT
GTTCAAACACAAAGAAGACACCTTGGAAACCGACGACCACGCCTTGCTCAACCTCTGCGCTGGGCCGAGGTCGTGCGTC
TGGGCCAGGAAGGCTGGGAGCTGGTGAGCGTTCAGCCACTCATGCGGGGCGTAACCGAGATCGGTAATCAAAACGCCCAA
GGCTGGGCTTGGGGCGTCGCTCTGCCGTCAGCTACCTGCTGTTTTCAAGCGCGCAACCTCATAA

Protein sequence: (SEQ ID NO: 134)

MSIKQPFYHVENIVI PYKTLTKGVAMFKHKEDTLEPDDHALLNPLRWAEEVVR LGQEGWELVSVQPLMRGVTEIGNQNAQ
GWAWGVALPVSYLLFFKRATS.

FIGURE 11

RL032

DNA sequence: (SEQ ID NO: 11)

ATGCTTAGAAACATCTCTATTGGAGTTTGTAGCCATGGCTGCTATGTTGGGCAGTTATGGGGTGGCTGCCGCTACATT
ACGATGCGGGTCGGCAATTGTTAGTGAGGGCGACTTGATTGATGATGTGCTTAGAAAAGTCCGGCAACCCTGATAGCCGTA
AAATTGAAGGGCCCCGAGTGGATGGTAGTGGCTATATAGTCCGGGGGGCTGCTACTGTCGAAAAGTGGGTATATGGACCA
AGGAATGGATGGTACCAGAAGCTTAGGTTTGTCTGATGGAAGACTAGTTCAGATAAAAGGCAGTATGGACTAG

Protein sequence: (SEQ ID NO: 135)

MLRNISIGVLLAMAAMLGSYGVAATLRCSAIVSEGLIDDLVRKCGNPDSRKIEGPAVDGSGYIVRGAATVENWVYGP
RNGWYQKLRFVDGRLVQIKGSMD

FIGURE 12

RLO33

DNA sequence: (SEQ ID NO: 12)

ATGAAACTTATCCTTGATTCGACGGACGCCTTCTAAATCCAAGCAACATGCTAGAGGCCCTATCAAAAGCAGGAAAAAA
TACAAGCATCAGCATAAGCAACGCGCAAGCATTAAATATAGAACTCTTCTCAAGGCAACAACCACTGCAGAAAAACAA
AAAACCTCTCAACAACCTTCAACGGCGCAGAGCTGACTGCTAACAACCTTCAGCAAGTCATAAACTCAGCAGGATCACTA
ACCAGAGTATCCACAATAGCCGCACAAGCCATTAAATATAAACACACTTCTTCCGCAATATCTACAGCAGGCAACTCAA
GAATTTTAGCGCAGAATTCAATGGAGCCCAACTCAGCAGCGACAACCTACTTAGAGCAGTAAATGCGGCAGGAACAAACA
CCAGCATAAGCGTCAATACCGCACAAGCGGCAATATAACCGCCCTTCTTCAAACCTATTCATGCAGCAGGTGACAAAA
ACATTCAGCGCAGAGTTCAATGGCGCTCAACTTACTTCAAACAACATTCAACAAGCTTTAGACGCGCAGGAACCCGAAC
ATCCATTAGCGTCAATACCGCACAGGCGGTTAATATAAGCACCCCTACTAGCCCTCATCAACTCTGCCAAAGACACGAAAA
AGTTTAGCGCGGACTTCAATGGTGCAAACTAACAGCAGACAACCTTCAGCAAGCGATCAGCGCTGCGGCCTCGGGTACC
AATATCAGCGTCAACACCGCTCAGGCGGCGAATATATCCACCCTTTTACAGGCCATCAACATCGCGGGCAACACTAAAA
ATTCAGCGCCAACCTTAAATGGTGCCCAACTCACTTCAAACAACATCCAGCAGGCGCTCCGAGCGACAGGATCAAACACAT
CAATCAGCATGAACTCCGCACAATCCGCCAACCAAGCACTCTACTTGAACCTCTAGACATAGCAAGTCCAGCAAGCAA
TTCCAAGCCAATTACAACGGTGGCATGTCTAATCCGAACAACCTACAACAGATAGTTTCCCGTGCAGGCGCCAGTACAA
CCGTGTTTATTTCCGACGCACAAGGCCCTACCAATCGCAAATATCCTTACCCTTATATCATCTGCCGGATGAGACTTATAG
CCGTGGATGAAAACACACCATCCACGGCTATACCCTAG

Protein sequence: (SEQ ID NO: 136)

MKLILDFDGRLLNPSNMLEALS KAGKNTSISISNAQALNIETLLKATTTAENTKNLSTTFNGAELTANNLQQVINSAGSL
TRVSTIAAQAININTLLSAISTAGNSKNFSAEFNGAQLSSDNLRAVNAAGTNTSISVNTAQAAINITALLQTIHAAGDTK
TFSAEFNGAQLTSNNIQALDAAGTRTSISVNTAQAVNISTLLALINSAKDTKKFSADFNGAQLTADNLQQAISAAASGT
NISVNTAQAAINITLLQAINIAGNTKKFSANFNGAQLTSNNIQALRATGSNTSISMNSAQSANQSTLLELLDIASSSQ
FQANYNGGMSNPNNLQQIVFPCRRQYNRVYFRTRPTNRKYPYPYIICRMRLIAVDENTPSTAIP.

FIGURE 13

RLO34

DNA sequence: (SEQ ID NO: 13)

```
GTGCAGTGGACTCACGAACAGTCACCGATCATCCAGTCGAAGGCACCGAAGATCCTGGTGCAGCCTTCGCAGGCACTGG  
CAAACTACCACCTGGTGGGCTTTGCCAGGTGCAACCTACCTGAGAATCCTCTATCTCTGCTACAACAGCTCGGTGG  
AGAAAGCCGCGAAGGGCAAGTTTCCCGCAACGTAGTGTGCAAGACCGCCACAGTCTGGCTCATGCGGTGTACGGCATT  
CAGTACGCGCCACAAGAAGACGAGAACCTGCGACTGACCGATATCGCCCGCGGACTCGATACCAAGACTGGGAGTTGGT  
ACGTGACGTGCTGGCCACGCTGAACAATACATGGCCAGCGCCGACGCGGAACCTCGGCCGACCGCACTTCCCGCGCTTCC  
GCGACAAGGCGTTCTCTACCACTGCTCAGGAACGCTTCTCAAGCAGGGCCTGGACATGGCGGAGTAGTCTGGAGGGCG  
ATGGTCGATCTCCAGGACACCGGCATGCTGATGCCCTTGACGGCTACCTGAAGCTGTATCAACTGAGCAAGCCGATTT  
GAGCCAGCGCTTCGACTGCATGCTCCTGGACGAGGGGACGGACATCAACCCAGTGATCGCGGACATTGCCATTGGCAGC  
GCATCAGAATGGCTATCGTCGGCGATCCCCATCAGCAGCTCTACCGGTTGAGGGGCGCAGAAGATGCCCTGAACAGCGAC  
TGGATGGCCGCGCGGAGGAGCACTACCTGACCCAGAGCTGGCGATTGCGCCCCGCGATCGCACAGTGGCCAAACATCAT  
CCTCTCTACAAGGGCGAAACACGGAACTTCAAGGACTGGGTCCGAGACGCTGGTGAAAAAGTCCCTCCCGCCGAGC  
TTCCTCACCGCACTTTCATTACCCGACCGTTATCGGCGTCATCGAGAATGCCCTGCAGCTGGTCCGCAATCATCCGAG  
CCCAAATTCACCTGGGTAGGCGGTATCGACAGTTACTCGCTGCGCGACCTGGAGGATCTGTACGCATTAGCCGAGGCT  
GCGCCAAAACGCTCCAGAACAAAGAACTGCTCCGTGACTACCGGACTACACCCAGTACGTGGAGATCGCCGAGATCAGCC  
AGGACGGTGAGATGCTTCGCTCGATCAAGATCATATCGACCTACCTGATCTGCCTGCGCGGATCCTTGAGCTTCGCTCA  
CTGACCCCTTGACGATGAGCTGGACGCAACAATCACCTGACCCCGCACACAAGGCCAAGGGGCTGGAATGGGATTTCTG  
TTGCTGTACGACGACTTCAACGCGGACCCGCTGGCCCCGACACCGACCCAGGCAAGCGCGACGATGAGTTGAACCTGA  
TCTAGCTCGCAGTGACCCGCGGATGAAGATCCTTGCCATCAACAGCCTGGTGCTGTCGATCATGCAGCGGTACGTGGAC  
GACAGAAAACCTGAAGGAGCAGATAGCTAGCTGTAATAAATGA
```

Protein sequence: (SEQ ID NO: 137)

```
VQWTHEQSPIIQSKAPKILVRAFAGTGKTTTLVGFARSNPTLRILYLCYNSSVEKAAKGKFPNNVCKTAHSLAHAVYGI  
QYAHKTKNLRLLTDIARGLDQDWELVRDVLATLNNYMASADAELGRPHFPRFRDKAFLTSAQERFLKQGLDMARVVWRR  
MVDLQDTGMLMPLDGYLKLYQLSKPDLSQRFDCMLLDEGQDINPVIADIAHWQIRMAIVGDPHQQLYRFRGAEDALNSD  
WMAGAEHYLTQSWRFGPAIAHVANIILSYKGETRKLQGLGPQTLVKKSLPPDLPHRTFIHRTVIGVIEALQLVRNHPE  
PKFHWVGIDSYSRLDLEDLYAFSRGLRQNVQNKLLRDYRDTQYVEIAEISQDGEMLSIKIISTYPLPARILELRS  
LTLDDELDTITLTAHKAAGLEWDFVCLYDDFNADPLAPDTPGKRDELNLIYVAVTRAMKILAINSLVLSIMQRYVD  
DRKLEQIASCKK.
```

FIGURE 14

RLO35

DNA sequence: (SEQ ID NO: 14)

```
ATGTTCCGGGTCGCTGATCGGCGCAATCATCGTGGAGTGGGTATGCCTGTATTTCTTCTGGCCTGACGCGGGCTGGAAGCA
TGCCCAGGCCATGTTTGAGTACGAACCTCAGTTGGCTGTGCGAGGGGCTGCTACACAGCGTCGTCGTGCAGGAGCCAGGTC
GAACCGCCACCTGGCTGGCCAGTTGGCCTATGACTGGTTGTTTCGTGAAGACCGGGATGGTCGACTGGATGACCAACATG
ACTACCATCGGCGAGGCCCGGCCACGGAGCCCGCTGGACGTTTCGCTATCTCACCGCCACGGTGCTCTCCACGCTGCAGAA
CTACGGCCTGGCCGCGCTGTACACGGTGTGACATTTCGTGCGCCTGGTGATCCTGGTCATGACGATCCCGTTATTTCG
TGATGGCCGCGTTTACCGGCCCTGGTGGACGGCCTGGTGCGCCGGGACCTGCGCAAGTTCGGCGCCGGCCGGGAGTCCAGC
TACCTCTACCACAAGGCGCGCGGAGCATCATTCGCTAGCGGTCGTCCCTTGGACGCTCTACCTGGCAATCCCCATCAG
CATCAATCCCTGCTCATCTGTTGCCCTGCGCCGCGCTGCTCGGCGTAGCGGTATGCATCACAGCATCCACCTTCAAAA
AGTACCTGTAG
```

Protein sequence: (SEQ ID NO: 138)

```
MFGSLIGAIIVEWVCLYFFWPDAGWKHAQAMFEYELSWLSQGLLHSVVVQEPGRATWLAQLAYDWLFVKTGMVDWMTNM
TTIAQARPRSPLDVRYLTAHGVSTLQNYGLAALYTVLTFVVRVLVMTIPLFVMAAFTGLVDGLVRRDLRKFAGRESS
VLYHKARGSIIPLAVVPWTLYLATPISINPLILLPCAALLGVAVCITASTFKKYL.
```

FIGURE 15

RL036

DNA sequence: (SEQ ID NO: 15)

```
ATGAAGTTGAAGAAATTTCTTACAGCCTTTTGATAGCGGTTTCTCCACTCCGAGTGTGCGCTCAAGCTGTCCGATGCT
CGGTGGCGCCTTGATGTTGTGCGTGTATGCAGCCTGATATTCAGTGTGAGCATGGTTTTAAACCATCAGGTGTCCCTCA
GTCGGCAAGCTATGAATGTGGCTATGTACGAAGCGCAGCTTTATTCGAGCAGCGGAGGCGTGTCTCAATCACTTGAGC
GGCAATGTCGTGCCCTTGGCCGCGGGTAGAGCGCTCGTCAACGAAGCGCGGAACAATGTGAGCATCTGCGCGTTGAGTGA
TGCGGCGGAGGTCTGCTATTGACCGCTCGCACGCTCGGTGATCTCGGGAAAAGCGGCTGGCACTGATGTATCTGGTCTG
ATACCGACAAAGGCCCTCTGGTTTACCGGCTTACCGCGATGGTAGGCCCTCGGCAGCGATATCCAGCAGGATAACCAA
GAGGTGTACCGAGCCTTGCTGGCGACTCGTGGCGCGCTGTTCACTGGGTGACTGACGGTGGTACCCCTCAACGGCTGTA
CCTTTTGAATCCTTAGGCGATGAGCCGGGCGAGGGGTGGCTAGGCCCTGGAGATTCTCGGCGAAGACCTCGATTGATGT
TGCGCCGGAATGATGCCGGAACCTACATGCTGCTGGATCAGCATGGGCGAGTGTCTCGTACGGACGCGAGGCGCTG
GGGAGCGGTGCGTGGCGGACGCTTTTGGTGGAGACGGCTTCGGTTTCATCGGTGCTGGCCACTGCGCGAGCATATGGT
GCTTTTCCAGCAGTGGGTCTTCGAGCTGGGATCTGATCTATCACATCGGTATCGGTGCGCTGTGCTGGCTCTGTGGC
TCCCTCTGTTACTTGCCTCTGCGTTGGCACTCGCAGTCGGCATCCTACTGCATTGGCTGGTGGCGAGCATCGAGCGACGC
TTGATAGACCCCGCAAAGCGAGCCTTGAAGCATTTGAAGGAGAGCGAAGCCTTTTCCCGTGCAGTTATCCAGGCGCGCC
CGTCGCGCTGTGCGTGTGCGTGTGCGTGTGCGGACGCGCAGTGGTCTGGAATAATCCCGAGGCGCGCAATGGCTGGTGATA
GCGAGGCGATTGCCACGACGCGCGGAGATGGATTTCCAGGCGTTCGCGAGGAGTGTGAAGTGTCTGGAGAAGAACTG
GAAACCGAGGCGAGGCTACATCTTCTCAATTACAGCCCAACCGCTATAACGGTGAAGACGATTGTTCTGCGCCTT
CAGTGAATCAGTGCACGCAAGCGGATGGAGGCGGAAGTGGCTCGCGCAAAATCCCTGGCGGATGCTGCCAATGAAGCCA
AGACGCTGTTTCTGCGCACCATGAGCCATGAATCCGCACACCTCTGTACGCGATGCTTGGCAGCTTGAGCTGCTGGG
CGTACCGAGCTGAGTCGCGCAGCAGGCGGTTACCTAAAGGCAATCCAGCATTCTCTGTCGACCTGCTGCAACTGATCAG
CGATGTGCTTGACGTATCAAGATAGAGGCGGCGCAACTGGACCTAGAGTGGTGAATTTCTCCCGCTGGAATTGACCG
AAGAGTCTGTCGAGTCGTTACCGGTGCGCGCAGGCTGCGGCTGCAAGTGTATACCTGCCTCTCTGCGGAGCTGCCG
CTGCGCATGCGGGGGCGCGCGCTCGATCCGCGAGATTCTCAACAACTGCTGAGCAACGCGGTGAAGTTCAACGACAA
TGGCTATGTCAACGTCCACCTGAAGGCGAGCGTGGTGCATGCCGAATGTGTGATGCTGACCTGGCAGGTCAACGATACCG
GCATGGGGATCAACGTGAGGATCAGCGCGCTGTTTCGAACCGTTCTACGATACGCGCTCCGAGCATCCGCTCGCA
GGCAGGGGCTCGGCTTGTGATCAGCGAGCGCTGGCGCAGCTAATGAATGGCAGTCTGAACTGGTCACTGAGCTGGG
GTTGGGCGAGCAGCTTAGCCTCAGGCTTCCGCTTGAGCGGATCGCGATGCAAGCTGAGCCGAGGACCTAGCCGGGTGCG
CCGTCCAAGTGTGGCGCTGTCCGCGACCTAACGGAATGCCTGTGTGGCTGGATCTCCCGCTGGGGTGAAGGGCCATG
GTCGCGACGCGGAGGTGCTGGACGAGGCGGACGCGACCTCGCTGCTGGTCAAAGTGTACTGCTGGAGGGGGCGCCGAT
GTTGCAAGCATGCGCAGGATGCGGGTGGAGCTTCCCTCAGGGTGATGAGAGCGCAGGCACAGGGCGCGACTGGC
TGCTCGGGCTCAACAACTGAACGCGCTGCATCGTGTCTGCGGCTGGCCCATGGGCGTCTCGCTGATCCTTCAGCGCG
CCGATACGCTGCTCCGTTGCGCAATCTAGGTCTCCGCGTCTAGTGGTGGAGGATAACCGGATCAACGAGTTGATCTT
GAGGGACAGATGGAAGCGCTGGGCTGCAGCGTGGAGCTGCTCTTCGATGGTCCGAGGCGTGTGCTGACTGCCAGACGG
CCTGCTTCGACGTGGTGTCAACGATATCAACATGCCGAACATGAACGGATACGAGCTAACCGCGAGCTACGGCGCCAA
GGGTTCCGCGAGCGGATCATCGGCGGACGGTGAACGCCATGCGTGAGGAGCGCGAGCGCTGCATGTCCGCGGGATGAA
CGATTGCTGCTCAAACCGGTGGATCTGAATGCCCTTCAGAACTGCTTGATTAATATTCTCAAGGTGGATCGATGA
```

Protein sequence: (SEQ ID NO: 139)

```
MKLKNFLQPFDSGFSTPSAALKLLRMLGGLMLCVLCSLIFSVSMVLNHQVLSRQAMNVAMYEAQLYFEQREALLNHLS
GNVVPAAAGRALVNEAPNNVSIPLSDGGRLLLTARTLGLDLREKRLALMYLVDTDKGPLYVYRLTADGRPSAAISSITIK
EVYRALLATPSAPVHWVDGGTPQRLLYFESLGEDEPGEGWLGLEIILGEDLDSMLRRNDAGNYMLLDQHGGQVVLATDAEAL
GSGASRTLLRGDGFIFGAGPLPQHMLVLFQHVGSSSWDLIYHIGIGRLLALLWLPPLLASALALAVGILLHWLVRSIERR
LIEPAKRRLEALKESEAFSRAVQAAPVALCVLRRADAADVLENPQARQWLGDSEAIHDAAPRWISQAFAGGVKCSGEEL
ETEAGLHLHLYNTPTRYNGEDVLFCAFSEISARKRMEAEARAKSLADAANEAKTLFLATMSHEIRTPLYGMLGTLELLG
RTELSRQQAGYLKAIQHSSSTLLQLISDVLDVSKI EAGQLDLECFEVSFPLETTEEVQSFTGAAQAKGLQLYTCLSAELP
LRMRGAAASIRQILNNLLSNAVKFTDNGYVNVHLKASVVDAECVMLTWQVNDTGMGINVEDQPRLFEPFYQIRRSEHPVA
GTGLGLSISQRLAQLMNGSLKLVSELGLGSSFSRLPLERIAMQAEQPDLAGCAVQVLAIPVRDLTECLCGWISRWGGRAM
VATPRSLDEADATSLLVKVLLEGA PMFEAWPGCRVELSPQGDMEPQAQGRDWLLGLNNLNLHRLALGLAHGRLADPSTP
PIRLAPLRNLGLRVLVVEDNAINQLILRDQMEALGCSVELLFDGREALLHCQTACFDVVLTDINMPNMNGYELTAE LRRO
GFRQPIIGATVNAMREERERCMSAGMNDCLVKPVDLNLALQNCLINILKVDR.
```


FIGURE 16

RL037

DNA sequence: (SEQ ID NO: 16)

ATGAGCTGGAATCCTATCGGGTGGTGGTGGTTCGAAGATCAGCCGTTTCAGCGCGAATACCTGCTCAACCTGTTTCGCGA
GCGCGGCGTGCAGTACCTGGTAGGTGCCGGCGACGGCGGGAGGCGTTGCGCTGCCTGAAGCAGGACAGGTTCCGACCTGA
TCCTCAGCGATCTGATGATGCCGGGCATGGATGGTATCCAAATGATCCTGCAACTGCCGTATCTCAAGCATCGTCCGAAG
CTGGCGCTGATGAGCTCCTCGTCGCGAGGATGATGCTCAGTGCCAGCGGGTCGCCCAGAGTCTCGGCTTGTCGGTAAT
CGACCTGTGCCCCAAGCCGACTCTGCCCCAAGGCCATCGGCCAACTTCTGGAACACCTGGAAAGATGCCCTCAGGCAGAAGC
TGGAGCCGGAACCGACGAGACTCCGCATGGGCGCACGGCGTTGCTGGATGCCCTGCATAACGAGCAACTGGTGACCTGG
TTCCAGGCTAAGAAATCCCTCCACACCGGGCGCATAGTCGGCGCCGAGGCGTTGATACGCTGGAGCCACCCGCGAGCATGG
CCTGTTGCTGCCAGCTGTTTCATGAGTGATGTCGACGCTACCGGTCTGCACGAGCGCTTCTCTGGCGCGTGTCTGAAC
AGACCTGAACGCCAGGAATCGTGGCGCAGGGCGGGTTACGAGATTCCGGTTTCGGTGAATCTGCCGCCGACCTGCTC
GATAACCAGGAACCTCCGGATCGACTCTATGAGTACGTCCGCGCTCGCGGGCTTGTTACCAGCTCACTATGTTTCGAGTT
GACCGAGAGCAGTGTCACAACTCTGTCAAGTAATACTATGTCAGGTGCTTGTGCTTGGCATGAAAGGGTTCCGATTGG
CCCAGGACGACTTTGGCCAGGGTTACAGCTCGTTCTATAACCTGGTACGACGCTTTACGGAGCTGAAGATCGACCGC
TCCCTAGTCCAGGGATGCGTAGAGGATAACGGCCTCAATGCAGCTGTCATCAGTTGTATTGAGTTGGGTACACCGCTGAA
TCTCGACGTGGTGGCCGAAGCGGTGGAGACCTGCGAGGAACTGAATCTTCTTCTGCTGCTTGGCTGCGACCGGGCGCAGG
GTTTCTGATTCTAAGGCAGTGTCTGCTCGTGAGTTTCGAGCGGCAGTTAAGGGAGGACGGCCCCAGCCTCCTTGTTTAA

Protein sequence: (SEQ ID NO: 140)

MSWSYRVLVVEDQPFQREYLLNLFRERGVQYLVGAGDAEALRCLKQDRFDLILSDLMPGMDGIQMILQLPYLKHREP
LALMSSSSQRMMLASRVAQSLGLSVIDLFPKPTLPKAIQQLLEHLERCLRQKLEPETDETDPHGRRTALLDALHNEQLVTW
FQAKKSLHTGRIVGAELIRWSHPQHGLLLPSCFMSDVSATGLHEALLWRVLEQTLNAQESWRRAGYEIPVSVNLPPHLL
DNQELPDRLYEYVGARGACTSSLCFELTESSVTTLSSNYAGACRLRMKGFLAQDDFGQGYSSFYNLVTTPFTELKIDR
SLVQGCVEDNGLNAAVISCIELGHRLLNLDVVAEGVETCEELNLLRRLGCDRAQGFLISKAVSAREFERQLREDGPSLLV.

FIGURE 17

RL038

DNA sequence: (SEQ ID NO: 17)

```
GTGAAGTCTGCTAGTGCCTTGGAGCACGACAACAACTTTTGTCTCAATGGACAACCTCTCGCAGAGCCTGAGCATCGG
CTTGATCTGTGTGGTGGTGTGACCGTATTGCTGTTTACAGCATCTGTACTGGTCTGGGAGATTGTTTTCAGGAGGAGG
AGGACAAAGTCTCCTTCCACTTACCCCGTATGATGGATGTTATACGGGAGCATGAGGTATTTCTTGGGCGCATCGCTCGA
AAAAGCGACAAGACCAACCCAGAAGTACGACTAGCAGTGGTGCCTTTCAGCGGCACCTGTTGGCAAAGGAAACGGATT
AGCGGTCTATGAGGGACGGGAGTTTCTTTGCTATGCCATTCTACTGGCTACCAAGCACGCGTTGAGCGCGGATTCTCT
CGGGAGATCCGTTTTGCTCGGTGATTGCTCGCCAATTTCTACGGAAGCTTTCGGAGTGTTCGCGCTATCCCGGCCA
CAGTTACTGATCTTTGATCTTCCGGCAGCACCCGCTGGCAGTGCCTGCGATTCCCTCCACAGCGCAGCGTGACAGGTT
GAGCGGAGCTATCCCATGATAGTTCGAGCGCATTTGCGCGGCTTTCGCAACCCGCGGCTGGGGAGGACGCTCAGCGTG
TCCATTGGATACGCGCTGATCGCTATCGCGACTCGCGCTGGAGATGTTGGGAGTCCCGGGTTGATCTGCCGGAACA
CTCTGGTGGCAGCAGAGCGAACCATCTGATCATCGCTGCGAGCTGCTTGATCTCAGGCGAATCAATGACTTCGAACA
GTTGGTTGAGCGCCCGCATTTCGATTCTGACAGCCTGGTATCGCCGGATGGCGAGGTATTGCTCGGCGCGGCCCTGCGA
CCGGCTGAGGGATGCGCTGAACCTCACCCGACAGGGGGTCCCGGTTCACTGCTCAGCAGCCTGAGAAGCGTTGGCTC
GCGGTCTACCGAACCAGTACGGCAATTTCTTTCGCCACTCCCGGTGGTGGTGGCAGGTCTGCTGCTGACCCCGCGCT
GCTCCTGGCGGCTGGCTCGGGATGCGTTCGGTACACAGCAGCGCTCGTCAACCCGCTGATCGGGCGCACCGGCAACTGG
TGGAGAGCGACACTTCAGCCGGACGCTGATACAGACCGCGCGGCTGGCTCTGGTGGTGTGACCCAGGATGACAGCAA
CTGGTGACCTGCAACCATGATAGTTCGCGCGGCTGGCGGGGCCACCGAGATCCTTGGGCTGACTTCCAAGTGAAGCT
TTTCGATGCGCGTGGGCGAGTACAGGAGACATCTGTATCCAGGTGGTGGGCGCTATTTCAGACCCGCTTCGCGGCGA
CCCGCTATGCCGGCAGCAGGCGGTACTGTGCTATTCAACGACATCAGGTCCTGCTGAGGCGGAGACCGCGCTGTCC
AATGCGAAGCAGCAGCGGATGCCGCCAGCCAGGCCAAGACCTGTTCTGGCCCGCATGAGCCATGAAATCCGTACTCC
CCTGTACGGTGTCTTGGCACCTGGAGTTGCTCGACCTGACCAACCTGAACGAGCGGCAACGCGCTACCTACGCCACA
TCCAGAGTTCGTCTCGCAGCTCATGCAACTGATTAGCGATGTGCTGGATGTCTCGAAGATCGAAGCGGGGAGATGGCT
CTGACCTTGGCCGCTTCAATCCGCTGGACCTAGTGGGGAGTGCTTGGCAACTTTGCCCGCAGCGCCATGGCCAGGA
CTGCGAGTTCTATGCTGATCGACACCGAAGTGCCGGCGCACTGATCGGTGACGTGACCGGATTTCGCCAGGTGCTCA
ATAACTTGGTGAATAACGCGCTGAAGTTCAACGATATCGGACGGGTGGTCTGCGCGTGAAGTTGCTCTCCCGCAATGAT
GGTTCGAGCCCTGTTGCGAGTGGCAGGTGCGCGACACCGGTATCGGTATCGCACACGAACAGCAGGAGCGCTTGTTCGAGGC
GTTCTACAGGTTTCGGGAGCGCACCATGCCGGCGGCGACGGGGCTAGGACTGTGCTGCTGCGCATCTGGCGGAAATGA
TGGCGGCTCACTGCGAATGGTTCAGCGAGACGGGCTCGGCAGCAGCTTCAGCCTGGTGTGCTCGAGTTGCCCGAGGACGAA
CAGTCCGGGCTGGCTTGGCGGCGGGGCTCTTGAATCCGCTTCCGTCCATGTGCGCTCGCCCGTGGCGGAGCTAGCCGA
CAGCGTAGGGGCGTGGCTGAAAGCTGGGGCTGCAAGGTACAGCAGCGCGAGGCGCGCCCTCCGAGCTGGAGACTGTG
TGCTTCTGGAGCTGCTGCCGATGGCGGCGGGGCTGCTTCTCGCCCTGGCCAGGCCCCGGGTGCGCGCTCCATGGAT
GGCGCTTGGCAGCGGAGCTGCGTGAGGACGGCTGGCGTGTGCGCTGCAACAACCTGGCGGGAATCGGCGAGGCTGGC
GCAGGCTCTGGGTGGCGATATCCCCAGCAACGCCGGCAATGCTGCGCCGCTCGGGGAGACTCGACCTGGAAGTGC
TGGTCCCGGAGGACAACCCAGTCAACAGGCGCTGCTTCGCGAGCAACTGGAAGAGCTGGGTGTGCGGTGAGCCTTGCC
GGCGATGGGCGGCGAGCCCTGACGTGTTTCGACAGTGGTTCGACCTCTGCTCAGCGAGCTCAACATGCCGAACAT
GACCGGCTACGAACTGACCCAGGCGCTGCGGAACGAGGCGAGACGCTGCGGATCATCGGCGTGACCGCAACGCTGCG
GAGAAGAGGCGAGCGCTGCCGGCAGTGGGAATGAACAGTTGGCTGGTGAAGCCGATCACTCTGCATACCTTGCATGAA
CTGCTCAGTGAAGTTCGCTCGCGCAGGTGCTGCTTCCCGCGCAAGCGCGAGACCTCGGCGCGCCCGCGCAGCTCGACGA
CGGTCTCTCACCCGAGGTGCCGAACGATGCGCGCGCTTTTCTTGAAGCATGGGCAAGGACCTGGAGGCGCGCCGCG
AAGGATTCGCGCGCAACGACCCGAAGGCTGCGAGGAGCTGATCGCATGCGATGGCCGCTCCGTGGCGGTGATGCGTGG
CGAACGCTGTGGTGTGTCAGGGCGCCGAGGAAGGCTGCTGGAGTTCGCGCTTGAATGTTCCGCGGTGGAGATTGG
CGAGGTGCTCGTTTATATCGAGCAGGCGCTGGAGTTTGTGAGAAAGACGGGCTGA
```

Protein sequence: (SEQ ID NO: 141)

```
VKSASALEHDKLLKWTLSQSLSIGLICVVVLTVLLFSICYWSLGRLFQEEEDKVSFHFTRMMDVIREHEVFLGRIAR
KSDKTTQKYDYDVVPLQRHLLAKENGLAVYEGREFSFAMPFLLATKHALSADSSGDPFSLGVLLANFYGSFWSVSAYPAP
QLLIIDLSSGSTRLAVPSIPSTAQRDRLSGSYPMIVERILARLRTRPVGEDAQRVHWIRADRYRDSALEMLGARVDLPET
LWWHDEPNHLLIAASLLDLRRINDFEQLVERPAFDSYSLVSPDGEVLLGAAPATGLRDLNLTRQGVAVQLLSQPENGWL
AVYRTDYGNFRHSRWLVAGLLLPALLAGWLGMRWYTSVVNPNVRAHRQLVESDTFSRTLIQTAPVALVVLTDQDDQ
LVTCNHLAAQWLGGPTEILGLTSNWKLFDFARQVPGDICIQVGGRYLQTAFAATRYAGTEAVLCVFNDITVHCEAETALS
NAKRAADAASQAKTLFLNARMSHEIRTPLYGVLTLELLDLTLNERQRAYLRTIQSSSATLMQLISDVLVDVKIEAGQMA
LTLAANPLDLVREVLGNFAASAMAKDLQFYACIDTEVPAQLIGDVTRIRQVLNVLNNAKLFKTDIGRVVLRVKLLSRND
GRALLQWQVADTGIGIAHEQQLRLEAFYQVSGAHAGGTGLGLSICWHLAEMMGHLMVSETGLGSSFSVLVLELPEDE
QSGLACRPGLLSACVHVRSPVRELADSVGAWLKAWGCKVSSGEAAPSELETCLVLELLPMAAGPASSFPWGPVRVMSMD
APCQPELRDGDWRVGLHNLGAGIQGALAALGGDIPEQTPANACARSGRLDLEVLVAEDNPNVQALLREQLLELGCVRSLA
GDGRQALQLFDSGRFDLLSDVNMNMTGYELTQALRERGETLPIIGVTANALREEGERCRAVGMNSWLKPKITLHTLHE
LLSEFARAGVVLPAQARDLGPQALDDGLSPQVPERMRALFLETMGKDLEAARQAIRNDPKGLQDLHRMAGSLAVMRA
RTLVMCQGAEEGLLESRLCESAVEIGEVLVHIEQALEFVRKTG.
```

FIGURE 18

RL039

DNA sequence: (SEQ ID NO: 18)

```
ATGCGTCCGGGGTCAATAGTTGGAATTAGAACACAAGAGAAGCCTATGAGTAAGCTCAAGATAGTACTGGCCGATGACCA  
TCCGATCGTGCATATGGGCGTATGCGCATGCTCGAGCGCGACGGTCGGTTCGAGGTGGTGGGCGAGGCCTCCACGCCCA  
GCGAAGTGGTCGAGGTGTGCCGGCAGAGCGAGCCGCATATCGCCATTACCGACTACAGCATGCCCGGGGACGAGCGCTAC  
GGCGATGGACTGAAACTGATCGACTACCTGTTGCGCAACTTTCCTCGTACTAAGGTGCTCATCTTCACCATGGTCGGCAA  
CCGCCTGATCCTCGACAGCCTCTACGATCACGGGGTGTCCGGCGTGGTGTGAAGAGCGGCGAACTCGACGAGCTGCTCT  
TGGCGCTCGACGTGGTGAAGCAGAACCGCGTCTACCGGGGCGCGAACATGCTCGACCCGACCAAGTGTCTGGCGAACCGC  
GACGAAGTGGAAAGCCGCTTCGCGCGCTTGTGATGAAGGAGTTCGAAGTACTCCGTCACCTTCGTTTCCGGCAGCAACGT  
CTGCGATATCGCACGGCTGCTGAAACGTAGCGTCAAGACCGTAAGCACGCAAGGCTCTCGGCGATGCGCAAGCTGGGAA  
TGAACAGCGACCAAGCCTTGATGACCTTCTGCGTGCATGCCAACTTGTTCCATTGA
```

Protein sequence: (SEQ ID NO: 142)

```
MRPGSIVGIRTQEKFMSKLIKIVLADDHPIVRMGVCDMLERDGRFEVVGEASTPSELVEVCROSEPHIAITDYSMPGDERY  
GDGLKLIDYLLRNFPRTKVLIFTMVGNRLLIDSLYDHGVSQVVLKSGELDELLALDVVKQNRVYRGANMLDPTSVLANR  
DEVESRFARLSMKFEVLRHFVSGSNVCDIARLLKRSVKTVSTQKVSAMRKLEVNSDQALMTFCVHANLFH.
```

FIGURE 19

RL040

DNA sequence: (SEQ ID NO: 19)

```
GTGTCCAGTAAGATCCTGCTGCAAGGGGCACTGCTCGGCCTAGCAATGCTGGCCGTGCTGGACGCCCGAGCCGGAGTCAC
CGCCGAGCGCACTCGGGCAATAATCGCCGAGGGGCACCGGAGACGTCGCTGCTGGTCAACCAGAATGCCTATCCGG
TCATAGTGCAGACCTGGATCGACGATGGCGCCCCGAACTCGACACCGCAGTCTGCCCCGCGCGCCGATCATGCCGCTACCG
CCGGTGTTCCGCCTCGAACCCGGACAGCAACGCAGCCTGCGCCTGCTGCGGACCGGCCAGGCGCTGCCAGGGGACCGCGA
ATCGCTGTACTGGTTGAACCTCTACGAAATCCCGCCGCAAGCCACCGGGCTGCTGGCCGAAGGACAGTCACGGCTGACCG
TTACACTGCGCACCCAGATGAAAGTCATCTACCGCCCTCGCCCTTGCCAGAGGTGCGGAAGAAGCGCCACACCAGCTC
AGGTTGAGCGCGGGGGCGAAACACTACAGATGGAGAACCTACTCCCTATTTTCATCAGCCTCGCCGGCGCCGAGCTTGG
CGGCCACACCCGCTGGCGGCGGCCGAACTGTTGCCCCCTTCTCCAGGCGCGTCTGGCGCTCCGCCAGGCGCTGCCCG
GCGGCCAGGCCGAGGTGCGCTTCAGCTGGATCGATGACGGCGGCAATCTCCAGCAGGGACGGAGCCTGCTTCACTGA
```

Protein sequence: (SEQ ID NO: 143)

```
VSSKILLQGALLGLAMLAVLDARAGVTAERTRAIIEAGHRETSLLLVNQAYPVIVQTWIDGAPNSTPQSARAPIMPLP
PVFRLEPGQQRSLRLLRTGQALPGDRESLYWLNLYEIPQATGLAEGQSRLTVTLRTQMKVYRPRPLARGAEEAPHQL
RFERRGETLQMENPTPYFISLAGAELGGHTRLAAAEILLPPFSRRVLALRQALPGGQAEVRFSWIDGGNLQQGRSLH
```

FIGURE 20

RL041

DNA sequence: (SEQ ID NO: 20)

ATGAAACATCCCTGCGCGTCTGCTCTGCTCTCGCGTGTCTCGCTCGTCTAGCTGGGCGACCTGCTACAAGGTCA
GGCGGTAGGCAACGCCACGACTACTCCAACCCAGTACGTCCGGTGAAGGCTCTGCCGCCACTGGGCGGAGCTT
CGCATCTGCTCAACGGTTCCTCGGTCTACCGAGCGTGATCAACGTTCAGCGACGCCAGTCTTCAGCCCGCAGGTAGCTT
ATCGCCAGCTTCGGTGGCGCGCTCAGCCAATACGGCGACAGCGCGGCTACGACCCAGAGCGCGTGTCTTCTTCGCTGTGC
TCCAGCAACGATGTGTAGCAGATGTTCTCCACCAATGCCGACGATCTCTACAGCGGCTGGTACCTAGGAGGCGACAGTG
CGGGCAACTCGATTGGCTGCAGTCCGCTTACTCGCACCCCTGGCCCAACGCTGCTGCGCCTAACCCACGTGGAAC
GGGCGATATTTACCGATGCTCGCGCGAGCGTCTGCTCGCGGGCTCGATCACTGCACTCGGAGGCTTTCACTGGTCAA
GGCGAAGAACCTCAGCGCGGTACGCGCCGAACGTTCGCGCTCCGCTGGAGTTCATCCGCTACTACTCGCCGACTACCG
CTTCGAGGTGTGATCGCCTACACCCAGCCCGCTGGTACATCGGCATCAAGGGCTCCGGGCTGGCCTACCCCAACGCTCGGC
CGCAGCCTAACGCCAACTACTCGGCTGGCACTACAATGGCGGGCGGCTACGGCCTGTACAACGACGATGACGTCAA
GCGCTATCCCACTGTTCCGTAACCAACGTCACGCCCCACGTTGTGTTCCCGTCGATTTCCTCAGTGAGATTAATGCCG
GCGCGAAGCTGAGATGCCCTTCGAGGTGGCCTTCAAGTCCCAACCGGAGATGATCAACAGCACCCTCCAGCGGTACT
GCATCGGCTGATCGGGCTTCAGCGCGGGCGCAGGCGCGCTCCGCTGCATCGGCTCGAGGAACGCAATGGCGGGCTCTC
CTACCTAGTTTCCGACCGGTACGGCGAGCCTGGTATGGCCCAAGGCGTGGGTATCCGCTTGTCTGCGCGACGGCAGTGCBA
TGAACCTGCTGTTAAGCGAGGATTTCCGCGATGGCGACGAATGCCGAACACCGGCTGTTATCGGATGATCGGCAACGCC
TCGAACAAGACTGCGCAAGCGGAGGCACTCAGCCAGTACACGAGACCTTCGGTGGCGCCTCGAAAACTCACCGTTGG
CAGCATGCGCCAGCGTTTACCCTGGGAGCGGGTGGAGGCGAGCGCGCGAGTGTGATTCTGTGTCAGTAA

Protein sequence: (SEQ ID NO: 144)

MKTSLRVLPLLLALLASSSWATCYKVTVAGNATTTSTNTQIRPGECSAGTWAGACDTCNGSLGLPSVINVDASFQPDGSL
TASSVAFSQYSGDSAGYDIPSRFFVCPAKEDVDVEMFSTNADLLYSGWYLGDGSAAGNSLTGLQSAYRTAWPNVLLRLTHVET
QGYFTDVWRRELTLGLLDIDSRGFLVPAKKNLSVAERAEFRAPLEFIYYSPDTSRLSYATQPAYGIAIKGPELAPYVNG
ASHNANYLGHWNWPGAIGLYNDVTLKRYPTCSVTNVTPHVVPFSPISLSEINAGANREMPFEVAFKCGTGVINSTASSGT
ALGIRASAGAQAASALGLRNLNAGLSYLVSDRYGQPGMAQQGVIRLLRDGSAAMNLLVSEDSAMGSAENATRGWYFVIGNA
SNKTGEAGGISQYSETFRARLEKLTGVSMLVTPGRVEASQGVVIRVQ

FIGURE 21

RL042

DNA sequence: (SEQ ID NO: 21)

```
ATGTTCTGCCACGTTGAGGCACGGCGCACCGGCAAACTGCCGCTGGCTCTGGGCGGCCTAGCGCTGGCCTTCGCGGGCCT
GGCCAATGGAGAGGGCGCAGTATCGTTCGACGACAGCTGTTGATGGGCTCGGGCCTCGCGGGCGGGACCCTAGAACGCT
TCAACCGGGCAACACAGGTGGACCCCGGAACCTACCATGTCTGATGTCTATCTCAACGGCAGCTACGCCAGTCGCACACAG
ATCGAGTTCGCCCCCGGGCGGGCGGCGTCAAACCTGCTTCGGCGAACGCTTCTTGGCGCGGACGCTGGGCGTCCGCC
CGCCTCTGAGGCGGGCGTGAAGCGCCTGGAGATTGCTTGGGGCTGGAAGAACGCTGCCAGGCTCGACCTTCAATCTCG
ACACGGCCTTCTGCGCCTCGATCTCTGGTGCCCCAGGCCCTGCTGGATATCAAGCCACGCGGCTACGTGGGTCCCGAC
GAGTGGGACCTGCGAGTAGCATGGGCTTCGTCAACTACGACGCCAGCTTCTATCGCTCGAGCTTCGACCGAGTAGGCGG
CAACGGCGACTCGGACTATGGCTACCTGGGGCTGAGCGGGGGCATCAATTTGGGCTGTGGCGCCTGCGCCACAGTCCA
ACTACAGTACTCCAGTATGCGGGAAACACCCGAGCGACTGGAACAGCATCCGCACCTATGCCAGCGCGCGGTGCCA
GGCCTGCGCAGCGAACTGACCTGGCGAGAGCTTACCGAGGGCAATCTGTTGGGAGCCTGGGTATCGCGCGGTGCG
CCTGGCCAGCGACGACCGCATGCTGGCAGACTCGCAACGCCGCTATGCTCCACAGGTACGCGGTACAGCGAACGCAACG
CACGGGTGGTATCAGCCAGAACGGCAAGAAGGTCCACGAATCCGCCGTGCTCCCGGTCCCTTCGTATCAACGACCTC
TATGGCACCGCCTACGACGGCGATCTGGATGTCCAAGTATTGAGGCGGACGGCAGCGTCTCGCGCTTTTCCGTGCCCTT
TTCCGCGGTTCCCGAATCGATGGGCCCGGCATCTCGGCTACAGCGCCACCCTCGGCCAAGCGCGCCAGTATGGCGACG
GCAACGACCTGTTGCGCGACTTCACCTATCAGCGCGGCTGACCAACTCGCTAACCGCCAACTCGGCTCGCGCTGGCC
GAGGACTATCTGGCGCTGCTCGGCGGAGGCGTGTGCTCGCCACGCCCTACGGAGCCTTCGGCTTCAACAGCATCTTTTCCA
TGCCACGGTGGAGAAGCGCCAGCGCAAGCAGGGCTGGCGTGTGCGTCTGAATACAGCCGGACCTTCCAGCCGACCCAGA
CCACTCTCACCTGGCTGGCTACCGCTATTCACCGAGGGCTATCGGACCTCGGCGACGCGCTTTCGGCGCGCCACGCC
GATGAGCACAACGACTCCTGGAATCCAGCAGCTACAGCAACGCAACCAAGTTACCCCTGCTGGTCAACGAGGCGCTGG
GGGTACGGCAACCTGTATCTGTCTGGAGCCACAGCGACTACTACGACGGCAAGAGCCGCGACACCCAGTTGCAGTTTCG
GCTACAGCAACACTTGGCGCCAGCTCAGTACAACTCGCCTATTTCGCGCCAGCAGACCACTGGTACCGCGATCTGAAC
GACGACTACGACCCGCTACTGCGCCGCAATAACAACCTGCGGCGGCGCGCTGCTGCAATCGCGCGGCTGACCTCGGCG
TTCCATGCCGCTGGGGTCTCCAGCCAGGCCCCGAATCTCAGCGCGATGGCCTCCCGCGCTTCGGCGCAGCGCGCGCA
GCAGTACCAAGACGGGCTCAACGGCACCTCGACGAAGACCGCAGCCTGAGCTACGCGATTGCCCGCGGGCGCGACAGC
GACAACACGGCAGCGATTTCACGGCAGTCTGCAGAAACAGACCTCGGTGGCGACGCTGAACGCGCGCTATGCCGAGAA
CAGCAGCTACCGGAGCTCAACACCGGCTGCGCGGCGCGCGCTGCTGCAATCGCGCGGCTGACCTCGGCGCGCTACG
TCGGCGACACTTTCGCCCTGGTTCGAGGCCAAGGGCGCCAGCGGAGCTGGCGTACGCGGTGGTTCAGGCGCGCGCTCAAC
GGCAATGGGTACGCGGTGGTGCCATCACTCTCGCCCTACCGCTACAACCCGTCAGCCTCGATCCGACGGCATGGGCGA
AGAGGCCGAGCTGCTGGAGACCGAGCGCAAGATCGCGCCATACGCGGCGCGCGCTGCTGTAAGTTCGCGCACTGA
CCGCTCACCAATTGCTAATCCAGGCCCAACTCGCGGACGGCAGCGCGCTACCGCTAGGGGCGCAATGTGCTCGACAGCCAG
GGTGTGAACATCGGCATGGTGGTCAAGGCGGCCAGGTCTATGCCCGCGCGAGGGCGCAAGGGCGCGCTGCGCGTGA
ATGGAGCGAACGCCAGGGGACGCTGTCTGCTGGATTACGACCTCGACACTGGCCCTCGCCAGGCTATCGAACCCGGAC
AGGCGGTGATCCGCTGACAGGGACCTGCACGCGCGTCTCGGAGGCACCATGA
```

Protein sequence: (SEQ ID NO: 145)

```
MFCHVEARRTGKPLALGGLALAFAGLANGAEQYRFDDSLLMGSLAGGTLERFNANQVDPGTYHVDVYVLYNGSYASRTR
IEFRPRAGGVKPCFGERFLRRTLGVRPASEAGVQAPGDCLGLEERLPGSTFNLDTALLRLDLSVPQALLDIKPRGYVGP
EWDAGSSMGFVNYDASFYRSSFDGVDGNGDSYGYLGLSGGINFGLWRLRHQSNYSYSSYAGNTRSDWNSIRTYAQRAP
GLRSELTLGESFTEGNLFGSLGYRGVRLASDDRLADSQRRYAPQVRGTANSNARVVISQNGKKVHESAVAPGPFVINDL
YGTAYDGDLDVQVIEADGSVSRFSVPFSAVPESMRPGISRYSATLQARQYGDGNDLFGDFTYQRLTNSLTANLGSRLA
EDYLALLGGGVLATPYGAFGFSIFSHATVENGQRKQGWVGLNYSRTFQPTQTTTLTLAGYRSTEGYRDLGDALSARHA
DEHNDSWNSSSYKQRNQFTLLVNQGLGGYGNLYLSGATSDYDYGKSRDTQLQFGYNTWRQLSYNLAYSRTQTTWYRDLN
DDYDPSLPPQYNLRHGSERSNTLTLSMPLGSSSQAPNLSAMASRRSGDSRGSSYTGLNGTLDEDRSLSYAIAAGRDS
DNHGSDFNGSLQKQTSVATLNAGYAENSSYRLNTGLRGAAVLHRRGLTLGPYVGDTFALVEAKGASGAGVRGGQGARVN
NGYAVVPSLSPYRYNPVSLDPQGMGEAELETERKIAPYAGA AVHVKFRTLTHGPLLIQAQLADGSALPLGANVLDSQ
GVNIGMVGQGGQVYARAEGDKRLRVQWSEPRGDACLLDYDLDTGPRQAI EPQAVIRLQGTCTPVSEAP
```

FIGURE 22

RL043

DNA sequence: (SEQ ID NO: 22)

ATGAATACTTTTCCACTGCCTCCGCTCCGTGCGGCTACGCTGGCGCTCGCCCTGCTGATACCCGCCATCCCGGCTCAAAG
CAGCGTGGTGATCATCGGTACTCGCGTGATTTATCCCGGCGACGCCCGGAAAGACCGTGCGAGATGATCAATCAGGACG
CATTCCTCCCAACGTGATCCAAGCCTGGATCGACAACGACGACCCCTCCTCCACCCCGGAGACTGCAAACGCGCCCTTTCTG
GTCAGCCCGAGCGGTGACGCGCATAGCCCCGGCAGCGGCCAGACCTGCGCCTCCTGTATACCGGGCTCCCGCTGCCCGA
GGATCGCGAATCGTTGTTCCATCTCAATGTGCTGCAGATCCGCCCCGCGACCTGGCCAAGGCCGAGCGCAACCAGATGC
TGCTGATGCAGCGCAGTCGACTGAAGCTGTTCTATCGCCCCGCGCGCTGCTTGGCGGCTCGGAGCAGCTAGTCGAGCAG
TTGCACTTCAGCCTGGTGCGAGCGAGCGGCAACTGGCGTGTGCGGGTGGACAACCCAGCGGCTACTACGCCTCCTTCGC
CGGCGCGATGCTGAGCATCGGCGAACGTGCTGGCGGCTGCTGTGAGCATGGTCCCGCCAAAGGCCAGGCCGAGTGGG
CGGCGGAACGCCCTTCGCCGCTCGCCCCAGGACCGGTCCAGTTGAACGCCCTCTTGATCAATGACTACGGCGCGCAATG
GAGGTCCAGCATGTTCTGCCACGTTGA

Protein sequence: (SEQ ID NO: 146)

MNTFPLPPLRAATLALALLIPAIQAQSSVVIIGTRVIYPGDAREKTVQMINQDAFPNVIQAWIDNDDPSSTPETANAPFL
VSPAVTRIPGSGQTLRLLYTGLPLPEDRESLFHLNVLQIPPRDLAKAERNQMLLMQRSRLKLFYRPAALLGGSEQLVEQ
LHFSLVQASGNWRVRVDNPSGYYSFAGAMLSIGERRWRLSSMVPPKQAEWAERPSPAPGPVQLNALLINDYGARM
EVQHVLPRL

FIGURE 23

RL044

DNA sequence: (SEQ ID NO: 23)

```
ATGAAACCTCAAAGTACTGCCCTGACTATCGCCGCAATTCTCGCATTGCCGGGTATCGCGGCGGCTGCCAATACCATCAC
CTTCCACGGAGAAGTGACCGACCAGACCTGTTCCGCCGTCGTGACGGACGAACCGACCCGACCGTGATACTCGACACCG
TACCGGTAAGCGCTCTTGACGGCGCAGTCGGCAAACCCGCCGGGAAACCAGCTTCACCCTGCAACTGACCGGTTGCGCC
GCTCCGGCGGCCGATGCCGAGGAGCACTTCAGCGTGATGTTCCAGGCGGTCAATCCGACCAGCGCCGGCAATCTGACCAA
TACCGCGTCCGCCGGCGCCACCGGCGTAGCGCTGCAGCTACTGACGGCACCGGGCGGCAGCGAGGTCAATCTGGCCGGCG
GGTCGGCCGTGGCTGCCGGTGACATCGTGCTCGCAGGAGGCGAGACCAGCACCAGCTACGACTATGCCGTCCGCTACATC
TCCGAAGCGACCACCGTCACTCCGGGACCGGTGCTCGGCTCGGTGACCTACACCTGCGTTACGAGTAA
```

Protein sequence: (SEQ ID NO: 147)

```
MKPQSTALTIAAFLALPGIAAAANTITFHGEVTDQTCSAVVDGRDPTVILDTVPVSALDGAVGKPAGETSFTLQLTGCA
APAADAEEHFSVMFQAVNPTSAGNLTNTASAGATGVALQLLTAPGGSEVNLAGGSAVAAGDIVLAGGETSTSVDYAVRYI
SEATTVTGPGVLSVTYTLRYE
```

FIGURE 24

RL045

DNA sequence: (SEQ ID NO: 24)

AGTCCGCACGGTAGTGACGACTGGAAGCGCTTCTGTGCTGCCAACCACTGGAGCCCAGCATGAGCCGGCGCGGCAATTG
TTGGGATATGCCGTGGCGGAATCCTTCTTCAGTAGTTTGAAGAAAGAGCGTATCCGCAAACGCATCTACAAAACCCGAGA
CATGGCCCGGGCGGATGTTTTGACTACATCGAGGTCTTCTACACCCGAACCCGGCGGCACAGTCATCTGGGTGGCGTCA
GTCCCGAGGCCTTTGAAAGCGCCTCG

FIGURE 25

RL046

DNA sequence: (SEQ ID NO: 25)

ATGGCTGAAGTCACTCAACGAGCAGAGCAGCAACAAGAGAGCCAGAAGACCCTTCTCGGCACCATCATCAGTACGCCCTT
CCAAATTTCTCGGCGTGATGTTTCGGGTCGCTGATCGGCGCAATCATCGTGGAGTGGGTTTGCCTGTATTTCTTCTGGCCTG
ACGCGGGCTGGAAGCATGCCAGGCCATGTTTGAGTACGAATCAGTTGGCTGTGCGAGGGGCTGTACACAGCGTCGTC
GTGCAGGAGCCAGGTCGAACCGCCACCTGGCTGGCCAGTTGGCCTATGACTGGTTGTTTCGTGAAGACCGGGATGGTCGA
CTGGATGACCAACATGACTACCATCGCGCAGGCCGGGCCACGGAGCCCGCTGGACGTTTCGCTATCTCACCGCCAGGGTG
TCTCCACGCTGCAGAACTACGGCCTGGCCGCGCTGTACACGGTGCTGACATTCGTGCGCCTGGTGATCCTGGTCATG
ACGATCCCGTTATTCGTGATGGCCGCGTTCACCGGCTGGTGGACGGCCTGGTGCGCCGGGACCTGCGCAAGTTCGGCGC
CGGCCGGGAGTCCAGCTACCTCTACCACAAGGCGCGCGGCAGCATCATTCGCTAGCGGTGTCCTTGGACGCTCTACC
TGGCAATTCCTCAACATCAATCCCCTGCTCATCCTGTTGCCCTGCGCCGCACTGCTCGGCGTAGCGGTGTGCATCACA
GCATCCACCTTCAAAAAGTACCTATAG

Protein sequence: (SEQ ID NO: 148)

MAEVTQRAEQQESQKTLTGIIISTPFQFLGVFMFSLIGAIIVEWVCLYFFWPDAGWKHAQAMFEYELSWLSQGLLSVV
VQEPGRATWLAQLAYDNLVFKTGMVDWMTNMTTIAQAGPRSPLDVRYLTAQGVSTLQNYGLAALYTVLTFVVRVLVILVM
TIPLFVMAAFTGLVDGLVRRDLRKFGAGRESSLYHKARGSIIPLAVVPWTLYLAIPININPLILLPCAALLGVAVCIT
ASTFKKYL.

FIGURE 26

RL047

DNA sequence: (SEQ ID NO: 26)

```
ATGGCTGGCCAGTACCCGTTGGAAGCGCTCTTGCGGCGCTGCCGTGGAGCTCTACACCACCACCGTGTGCTTCACCGCAGC
CGCGCTCTGCATCGTCGCGCGTGGACGTTCTCCCTCACTCCGCTGTTCCGGCATCGTGGCCGCGCTGTGCTTCGCCTGGC
TGGGTATCGTGCCTGGAAGCAGGCCGCGCTGGTGTCTCCGCTACCGCGGGAACATTCGCCGACTGCCGAAGTACACGATG
ACCAGCGCCGAGATGCCGGTCAGCAACGAACACCTGTTTCATCGGTAAAGGATTTTCGCTGGACGCAGAACATACGCAGCG
CCTGGCAGATACCTACCTGCCCCAGTTCCGCTCTTACGTGAGCCCTCGCCCTCTACGAGCGCGCGCGCGGTTGGAGA
AGCAGCTCGAGTTCGCCCCCTTCCCCCTGAAGCTGGTCGCCAAAGCCACTGCCTGGGACGTGGCTGGAAACCCCGCACGG
CCGCTGCCGCCCCGTGGCGGTTTGCTCGCTCCATGGCATCGAGCCGCGGGAACAGGACGTAGGCTGCAACTGGGCGA
GCGCGTCGGCCACACTGGTACTCGGCACACGCGGGTGGGTAAAGCGCGCTCGCCGAGCTGTTTCATCACCAGGATA
TTCGCCGCACTCACTGCCGGGTACGACGCCGCGGGTGAAGATGGGCGCGGACCCAGACGGTTCACACGGCTATCGG
CGCCGCGCGCAGAGGAGCAGCCGACTACGAGGTGGTATCGTCTTCGACCCGAAAGGCGACGCTGACCTGCTGAAGCG
TATGTACGTGGAATGCGAACCTGCGGCGCGCTGGACGAGTTCTACGTGTTCCACCTCGGTATCCTGACCTGTCCGGCAC
GCTACAACGCGCTCGGCCGTTCCGTCGGATCTCCGAGGTCCGACCCCGCTCGCCGCGCAGCTCTCCGCGGAGGGCAAC
AGCGCGCGCTTCGCGAGTTCCGCTGGCGGTTCTGTAACATCATCGCCCGCGCGTGCACGCGCTGGGTATCCGCGCTGA
CTACCAGCAGATCCTCCGGCAGTCTGTAACATCGATGCGTTGTTTCGTCGAATATGCCAGAAATACATCAGCGAGCAG
ATCCAGGCGCTGGGACACCATCATCCAGATCGAGGGCAAGCTCAACGACAAGAACATCCGTTCAACATGAAGGACCG
CCCCTGCGGGTCGTAGCCATCGACCACTGACACAGAAACGCATCGCCGACCCGGTTCATGGAAGGCTTGAAGAGCGC
CGTGCCTACGACAAGACCTACTTCCGACAAGATCGTGGCTCGCTGCTGCGCTACTGGAGAACTCACTACCGGGCGGA
TCTCGGAGCTTCTTTCCGCCAACTACGCGGACCTCAACGATCCGCGGCGGATCTTCGACTGGATGCAGGTTCATCCGAAA
CGCGCGTGGTCTACGTCCGCTCGACGCACTATCGGATACCGAGGTCCGCGCGCGGTTGGGCAACTCCATGTTTCAGCGA
CCTGCTCTCGGTAGCGGTCACATCTACAAGCATGGTGTGATGACGCGCTCGCCGGCTCGCTCGCCAGCGGCAAGTCC
GCATCAACCTGCATGCCGACGAGTTCAACGAGCTGATTGGCGACGAGTTTCATCCCATGGTCAACAAAGCGGGCGCGCC
GGCGTGCAAGTGACGGCTACACCCAGACCATGAGCGACATCGAGGCCAAGATCGGCTCCCGCGCAAGGCCGCTCAGAT
CATCGGCAACTTCAACAACCTGTTTCATGCTGCGGGTGCAGGACGCCACGCGCGAAGTCTTACCAATCAGCTCCCCA
AGGTCAGATCTACACAGCAGCGCGCGGCGGCGCCACGACGCGATCAACAACAAGAGGTAGCCTTCACTCC
AGCTCGCAGCAGGAGTGAGATGACAGCGTCCGATGCTCGAGCCGCGCCACATATTGGTCTGCCCAAAGGACAAGC
GTTCCGCTACTCGAGGCGCGCAATCTCTGGAAGATCCGAATCCGCTGCGCGCGGTCGCCCCGACGAGGTGATGCCGA
AAAGCTGCAGGAGCTGGCTGCGGTATGCGCAAGGGCCAGGCCCAACAGCGAGTGGTGGGAGGCGCGGATACTCC
GCCCCTGCAGGATGGTCTGCCCCAGGACCTGGTCGACGATTTCCGTCACTCGGCACCGGTGAGGATGCCGCTGA
```

Protein sequence: (SEQ ID NO: 149)

```
MAGQYPLEALLRPAVELYTTTVCFTAAALCIVAPWTFSLTPLFGIVAALCFAWLGIVRLKQAGVVLRYRRNIRRLPKYTM
TSAEMPVSNEHLFIGKFRWTQKHTQRLADTYLPQFASYVEPSPLYERARRLEKQLEFAPFPLKLVAKATAWDVAWNPAR
PLPPVGGPLRLHGIIEPREQDVGLQLGERVGH TLVLGTTTRVGKTRLAELFITQDIRRTHCRVRRRRVVKMGRRTQTVHHGYR
RRRAEEQPDYEVVIVFDPKGDADLLKRMVVEECERAGRLDEFYVPHLGHDPDLSARYNAVGRFGRISEVATRVAGQLSGEGN
SAAFREFAWRFVNI IARALHALGIRPDYQQLRHVVNIDALFVEYAQKYISEHDPRAWDTIIQIEGKLNKNI PFNMKGR
PLRVVAIDQYLTQKRIADPVMGLKSAVRYDKTYFDKIVASLLPLEKLTTRISELLSPNYADLNDPRPIFDWMQVIRK
RAVVYVGLDALSDTEVAAGVNSMFSDLVSVAGHIYKHGVDDGLPGSLASGKVRINLHAEFNEIGDEFIPMVNKAGGA
GVQVTAYTQTMDSIEAKIGSRAGQIIIGNFNLFMLRVRETATAELLTNQLPKVQIYTTSTPASGANDAINNNKKVAFPS
SSHDQVQMTSVPMLEPAHII GLPKGQAFALLEGNLWKIRMP LPAVAPDEVMPKSLQELAAAGMRKQAANSEWNEAPGYS
ALQDGLPQDLVDDFRHLGTGEDAA
```

FIGURE 27

RL048

DNA sequence: (SEQ ID NO: 27)

ATGACTACTCATCTGATCACCCCTAGTCATCAAGCAGCCGAGCGACGCTCAGGCACGCCAACTCATGTACCAGGAGTTGCT
CGGACTGATCTCACGCTACGGCGGTGAGGTGAGGTCCAAGGCCTTGGAGGACGAGTCGACCCTCTGCGAGCTGCTGGTGC
AGATGCTGCCTGATCATGAGGTAGAGCAAGCCAGGAAACAGGTGCTCGAACTTCATGCCAAGGGCCGCCTGCAGGCGCCG
GCAAGCCTGAAGGTGTAA

Protein sequence: (SEQ ID NO: 150)

MTTHLITLVIKQPSDAQARQLMYQELLGLISRYGGEVTSKALEDESTLCELLVQMLPDHEVEQARKQVLELHAKGRLQAP
ASLKV

FIGURE 28

RL049

DNA sequence: (SEQ ID NO: 28)

ATGAAGAAGTTCCTTGCCACGCTGGCATTGTCACGGCGTTTCGCGACTCAAGCCTGGGCCGCGGGCTGATCGTTGTCTGA
AGACCTCGGCGGCGCTCGGCGCTCCCCTACTACCAGGGCCTGGATCCGCAGCCATCCGCTTCCGCACCAGGACCTGGCG
ACCTGGGCGTCCGTGGCTCAGGTGCGTTTCCAGTTCGCTCCGCCCGCCTATCGCCAGGACGGGTCCAGGGGCGCGCCATC
AACGCTCCAGGCCTGCAACTGCTGTTCTGCTGGTCGGCGACGACACGCTGTCTCGAACCTGGCTGAAAGAGCGAGGCGACGA
GCTTCGAGACCTCCAAGCCGTGGGCTGGCAGTGAACGTGGCCAGCGAAGCGCGCCTGACGGAAATCCGGGCTGGGGGA
AAGGACTTCAGATATTGCCGCGCGCGGCGGACGACCTGGTCGACCGGCTAGGGCTGCAGCATTACCCCGCCCTCATCACA
TCCACCGCCATCCAGCAGTAG

Protein sequence: (SEQ ID NO: 151)

MKKFLATLAFCTAFATQAWAAGLI VVEDLGASALPYQGLDPQPSASAPGPGDLGVRGSGAFPVRSARLSPGRVQGRAI
NAPGLQLFLVGDDTLRSRWLKERGDELRLQAVGLAVNVASEARL TEIRAWGKGLQILPAPADDLVDRGLQLHYPALIT
STAIQQ.

FIGURE 29

RL050

DNA sequence (SEQ ID NO: 29)

```
ATGGCAACGTCTGTAGTTCGAGCCCTCCAGTTGGCCACCCTGCTGGTCCTGGTCAACATCGCTCAGGCCGCCGTGGATCC
ACCGCCGGCGTACAAGCAAATCGCCCTGCCCAAAGGGGTTCGGCCGAGGTGCTCTACTCGGTCGCGCTGACCGAGAGCA
AGGTCTGTGTCGCGCGCAATACGTTCCCTGGCCCTGGACATTGAACGTCGCCGGGAAATCTTACTACTACGCGACCCGC
ACCGCCGCCTGCACAGCGCTACTCGCGGCGATCAACCTCTACGGGGCCAAGAGCGTCGATTCCGGCCTCGGCCAGGTCAA
CATCGGCTGGAACGGACATCGTTTCTCCAGCCCCTGCGAGTCCCTGGATCCGTACAAGAACCTGGACGCCACCTCCGACA
TCCTGATCGAGCAGCGGGACGCCCTGTATGCATCCGCCCCGGGAAGACCGGTGGACTGGATCCAAGTTGCCGGCCGCTAC
CACCGCCCGCGCGCGCGCCTGCCGCCAAATACCGTAGGACGGTTTCCCGCCACCTTAGCCAAGTTCTCGGCGTCAA
CCTACTGGTGACCAATCCATGA
```

Protein sequence (SEQ ID NO: 152)

```
MATSVVRALQLATLLVLVNIAQAAVDPPPAYKQIALPKGVPAEVLVSVALTESKVLLRGEYVWPWTLNVAGKSYYYATR
TAACTALLAAINLYGAKSVDSGLGQVNIWNHGRFSSPCESLDPYKNLDATSDILIEQRDALYASAPGRFPVDWIQVAGRY
HRPAGGAPAAKYRRRTVSRHLSQVLGVNLLVTNP
```


FIGURE 30A

RL051

DNA sequence: (SEQ ID NO: 30)

ATGATCAGAACCGTATCGCTCCTGTCCGGCCTGATGCTGCTGCTGAGCTATCCCGCAGCCGGCCAGGAGGCGGCGGCAAG
CCGAGAGGCCAGCAGCCAACTGTCCGGTAGCCAACTCGGCACGCTGAAACAACAGACATCTCAGAGCGACCTGGCCCGAGG
AGTGGGGACTGAACCAACAGGAATGGACCCGCTACCAGACGCTCATGCAAGGCCCGGGGCGCTTACTCGCCTGCTATT
GATCCGCTGACCGCGCTGGGCATCGAGGCGCGATCGGCAGAGGAACGGCGGCGGTATGCCGATCTACAGGTCCAGGCCGA
ACGGCGCCGGGTCGAGAAGGAACCTCGCTACCAGCGCGCATACGACGAAGCCTTCGCCCGCGCTATCCAGGCGAGGGGG
TGATCCGCTCACCAGAAAGCAGCACAGCCAAACCCGTCGGGCACGCGGAACATGAGCCCGAGCGTTGAGAGCAGCGGGCG
CTGGCCCTGTTCTGTCAGGACAACTGCACCGCTGCATCCAGCGGGTCCGCGACCTGCAACATGCAGAAAAGGAGTTTGA
CCTCTACTTCGTGCTAGCCAGAACGCGCAGAGCGAGTGCAGCGCTGGGCAATCCTCGCCGGCATCGACCCGAAGAAGG
TTGCGAGCAAGCAGATCAGCTCAATCATGACGAGGGCCGCTGGATGGCCCTAGGACTGGGCGGAGCCCTTCCCGCCCTG
GTACAGGAGGTGAACGGCCGATGGCAACGTCTGTAG

Protein sequence: (SEQ ID NO: 153)

MIRTVSLLSGLMLLLSYPAAGQEAASREASSQLSGSQLGLTKQOTSQSDLAQEWGLNQEWTRYQTLMOGPRGAYSPGI
DPLTALGIEARSAEERRRYADLQVQAERRRVEKELAYQRAYDEAFARAYPGEVIRLTESSTANPSGTPNMSPALQSSGR
LALFVQDNCTACTQVRVDLQHAKEFDLYFVGSQNDAAERVRRAILAGIDPKKVRSKQITLNHDEGRWMALGLGALPAL
VQEVNGRWQRL.

RL052

DNA sequence: (SEQ ID NO: 31)

ATGAAACGCCCATCCCTGCATCAATGATTCTTGGCCTCTGTTTGACGGCAATGGCCGGCCTGCTGAGCTACCAGCAGTA
CCAACTCGTTAGCTCCGATCAGGCGTGGACAGTGCCTCGGAAAGGCCCTCGCTGGAGGCGATCCTGGCTCGTTGAGTC
GAGTCGACGAGCGCCTCGACGCGGTGGATGGACAGCACCTGGTCAGCAACGAGGACTTCGTTTACGGCCAGCAGGCGCTG
TCCAACCGAATTGACGCTGCGCAGGCGTTCCGCAAGCAGGCTCCGATGCCGTCGAGAACCTGGCTCAGACCACCGCCTC
GGCCGGCGACCTCTTGGTGCTCAAGGCAACCGTGGAGACACTGGACGGTTCTGTCCGACGCTTCAAGAAAAGCAGGCCA
AGGCGCCGCGCTGATCGTGCAGCGCCAAACGCCCATACCCGCAAGCCCAAGCCGAAACCCAAACCGATGGAGCCC
CCGCCCTTCTCGATCCTTGGCGTGGAGTATCGCGGGGAGAACGGTTTCTGTGCGTTGCACCTCCGGGATCCACCCAGCT
CAGCCAGATCTACCTCATTCCGCCGGGAGATGCCGTGCGCGGACGACCTGGCGACTGACCGACCTTGACGATGGTACCG
CGCACTTCGACGTGCGCGGCACCTCGCGCAGCGTTTCGCATCCAACCATAG

Protein sequence: (SEQ ID NO: 154)

MKRPRSPASMILGLCLTAMAGLLSYQYQLVQLRSGVDSAAEKASLEAILARLSRVDERLDAVDGQHLVSNEDFRSGQOAL
SNRIDAAQAFKQASDAVENLAQTASAGDLLVLKATVETLDGSRVTLQEKQAKAPLIVPAKRPPIPAKPKPKPKPMEP
PPFSILGVEYRGERFLSVAPPGSTQLSIIYLIRRGDAVAGTTWRLTDLDDGTAHFDVAGTSRSVRIQP.

RL053

DNA sequence: (SEQ ID NO: 32)

ATGCCGCGCCGCTTGATCCTCTCGGTACGGAGCGGGATATCCTATTTGCACTGCCGGTAAGCCGAATGACCTCACTGAC
TACTCCACCTCAACGAGTCCGCCCCATCGTCGATCCGCCAGCGACGCGGCGATGCCAATCACTTGGTTTTTTCGGTGCAG
GTCAGCCTGCTGTCTATCCAGCGTTCAGCCCTGATGCGCGACGAAGAGCCCCCGAG

RL054

DNA sequence: (SEQ ID NO: 33)

ATGGCCGAAGCTATCAGAAAGGATGCAATGATGACAAAACCTCTACTTTGATCTTCTGAACTCGCCTGCCGAGGCTCATTC
GTGCATACAAAAGTCTTTATCTGTGCAGGCAATCTCCACAACCTGTCCCAATACTGGAGTTTCTTCCGAAACCGTATACG
CCTATGCATCGTACATAAATGCATTAAGTATCGGTCAACGCATAGATCCTGCATTACCCAGAGCTTAACGAGTGCCATA
TCCAACCTGGCAGGTCGCCCATTGTCAGTAAGCGACATTTACCAAAAAATTCATGAAACCACTGAGAACACCTGTTGA
GATGGGCGTTCTGCTAATAGCATCACCTTTGAGGAGTATCAGGCCACCATAAATCAGCAAGCCATCAACATGGTTCAAG
ATATGCAGGATGGAGACAAAGGTGAGAAGGTGGAGGCCCTCCAGGCCAATATGAGTTTCTGTATGGACAGGAGATAAAT
ACTGATTTTCATCGCTCGTAATGAACTCGCTGCTGGGCGAGAGCGAAACCGTCGCAATAGTTTACGGGGCATATCACCAT
CGGTTACGGCTTCGATACCTTCGTGCATGAAGCGTCCGAGCTAAACTCTTGAATCTTGTGGTTCTACGCGACAGAAGG
TATTACCTGCATTGCAGCTATCAACGTCCGACCCAGGCTTCTGGAGCGTCTATGCCTTGCTGGGACAAAGTCTCACGGAT
GACGATGGGCTATTACTCTTTAGTGCCAAAGCGCGAGCTGTTGTTCAACGCATAGCAAGCAACAGTTTGCAGGTAAGTG
GAATGGGCTACCCCGAGCTATCAAAACGGTTGCGCTTGATCTATATTATCAATATGGGCAGACTGGTAATTTTCCAAAAT
TTCAACAAGCTATAAATAGCTATGTTGGCCGCGAGTCATCCATGAACTTAGAACTGGAATGGTGTACCGGATGATCCT
CTCCAGTTTATTACAAAACGATTGGAAGAGCGAGCCAAGTATCTGGCAATATCCTTCAACTATGAGCAATGA

Protein sequence: (SEQ ID NO: 155)

MAEAIKRDAMMTKLYFDLLNSPAEAHSSIQKSLSVQAISTTVPILEFPSETVYAYASYINALSIGQRIDPAFTQSLTSAI
SNLAGRPPIAVSDIYQKIHETTLRTPVEMGVRPNSTTFEEYQATINQQAINMVQDMQDGDGKGEKVEALQANMQFLYQGEIN
TDFIARNELAAGQRAKTVAIVQGHITIGYGFDTFVHEASELNSLNLVGSTRQKVLPAQLSTSDPGFWSVYALLGQSLTD
DDGLLLFSAKARAVQRIASNQFAGKWNGLPPAIKTVALDLYYQYQGTFGNFPKQAINSHDWPVIELRNWNGVPNDP
LQFITKRLEERAKYLAISFNYEQ.

FIGURE 30B

RL055

DNA sequence: (SEQ ID NO: 34)

ATGAACAACACAGTGAGCGAAACGCAACAGATCAATATTACCAAAATCCGGGGCAGTCTATTTCGGGTCTCTACAAGGG
GCTGGCTAACAGTGTCTCTGGCCAGCCATTTCAGAGGTACAGCTTGTGGAGGCTTGGGATATCCCTCTCGTACTCC
ATCCGGAGTTTGTGCCTAACGGAGATGTCTCGAAAATCGATAAGGAGTACGGAACGATCCTTGCTGCTGAGTCAGCTCAG
GTTATCTGCTTCAACTCCAAATGGCTCAAGACAAGGCTAAGGCGTGCAGGGAGGTTACAGCCTTGATCAGTTCTGTCTC
CTCCAATCTCAATACCATTAAAGAGTCGTATGTTGCTAATTATCTAAACCTGCTGAAACAATCACCAGAACCGATACCCGA
CTAGCGTCGGAGTTGAGATCATGTGAGGTGGCAGTCCGAACAGGATTCTGGAATCGAGGTCTCTTACGGTGCCAGTCTC
GGCGTCTAACTCAATCAACAATTCAGGCGATGAATCTGCCTGCCAGTCTCAAACAGTTGCTCACTCAGGGAATCGGTGT
GAAGCTTTCTCAGCCTGAATATTGGCCTGCTTACAACAACATAGCCACTGGTATTCTGTTATACAACCGGAGTGGCGATAA
CGTTGGCCTATTGGGCCACGGTTTAG

Protein sequence: (SEQ ID NO: 156)

MNNTVSETQQINIYQNPGQSI SGLYKGLANQCS PGQPFPEVQLVEAWDI PLVLHPEFVPNGDVSKIDKEYGTILAAESAQ
VILLQLQMAQDKAKACGEVTALISSVSSNLNTIKSRHGANYLNLLKQSPNRYPTSVGVEIMSGGSPNQDSGIEVSYGASL
GRLTQSQQLAMNLPASLKQLLTQIGVKLSQPEYWPAYNNIATGIRYTTGVAITLAYWATV

RL056

DNA sequence: (SEQ ID NO: 35)

ATGACCCAAGCTGCGAAAATACCAGCAAATGAGTACTCATTGGGGGATGGAAGAGGCTACATCAATATCTGGCCGAAAA
GGATGAGGCTCAGGCATTTCTTATCCATAATGATGGGCTTAATGGGGCTACATGCAGCCTTAAAGGCACTCTTAGAGATA
ATAAAGGAGTGGTGCAATTCGCCGTATTCTCTGCTTCAATGTTGCTAAGTATCACCCAGACAGGGCTGCTGTAGTAAGC
GTCAAACGTGAGGAAAATTCCCAAGCTGCTCTGTCATGTTGCGGTCTAGAGTTTGGTTTGAAGGAGCCTATAGCGTCCC
GCCAAGGGCTGTCTATATGCAAATAAGGAAAAAATCGACAAATGTTGGGTATGATTGAGAAAAAGAGCTTGATG
CCGCTCGCGCCTTATCAAATAAGCTTTTGTGAGACTGCGCAACCGAGCTAGCCTATCCTGCCAAGATATACTTGACGAAC
ACACTTGCCATGATCAGTGCTGAAAAGGGAGAGAATGCTCGCTGTTTGGAGTATGCCCATCGGGTGCAAAAGCAAATTCC
TGTAAGAGATGACGGCAACCGGCTGAAGACTTGCTCCCGCGGAGCAGCGTTTCTGCTATGGAACAACCGCCCAAGGCTG
ATGCTCTGTCTGAGCGATGCAGCGACGAGAAATAA

Protein sequence: (SEQ ID NO: 157)

MTQAAKIPANEYSLGDRGYINIWPEKDEAQAFLIHNDGPNGATCSLKGTLRDNKGVVHSPYSSASCLLSITQTGLLSVS
VKREENSPSCSAWCGPRVWFEGAYSVPFKGCYYMQIRKKTRQMLGMIEKKELDAARALSNKLLSDCATELAYPAKIYLTN
TLAMI SAEKGENARCLEYAHRVQKQIPVRDDGQPAEDLLPAEHAFAMEQRAKADALSERC SDEK

RL057

DNA sequence: (SEQ ID NO: 36)

GTGCTGGTAGAGCGTTTGGCCGACTGATGTTGAATTCGCGGGCGAGCTGAGCCTTGGGCTCGCCGCGCGCTGCCCGCAGCC
CCAGGGTAGCACCTGCTTGTGCGACAAGGCCTCTTTCGGGCCCCGGTACGCGCAGAGCTTGATATCCTCGCGCTACCGCG
CTGGTGCTGCTGTATGCTGCTCAGCAAGCCAGCGCGCGGTCTTTTCAGGGTAAGCGTACGGCCAATACACCTTTACTTA
GGTTGA

Protein sequence: (SEQ ID NO: 158)

VLVERLPDVEFAGELSLGLAGRCPPQGSTCLSDKASLRPRYAQSLISSRYRAGAACMLLSKPAAGLFRVSVRPIHLYL
G

RL058

DNA sequence: (SEQ ID NO: 37)

ATGGATATTGCGCTGGAGATTTAGCGCTTGAACAGCTGTTGCTAGAGCCGAATCGAGAAAGAATGATCGACTGCTTAA
ACAGCTGCTTACCGAAGACTTCGTTGAATTTGGAGCTATCGGCAAAAGCTGGACGAAAGCGGAGGTGATCGTGGGACTAA
AATCCCAGACTTGGATCAAAAGGACAATCGAGGATTTCAAACCTCGGTGTGCTTGAGATGGTGTGCGGTTAGCAACGTAC
CGATGCCGTCATCAAAATGCTAATGGCGATGAGTCGTTATCAATGCTAGCTCTGTTTGGAAAACCTACGAAGATGGTTG
GCACATGGTGTTCACCAAGGCACGAGGGTCTCCGAGTAG

Protein sequence: (SEQ ID NO: 159)

MDIRLEILALEQLLEPESRKNDRLKQLLTEDFVEFGAIGKSWTKAEVIVGLKSQTWIKRTIEDFKLRVLADGVALATY
RCRHQNANGDESLSMRSSVWKTYEDGWHMVFHQGRVSE.

FIGURE 30C

RL059

DNA sequence: (SEQ ID NO: 38)

ATGACTTCCTCGCCCAACCTTGACCAGATGACCCCGGAACAGCTTCGTGCCTTGGCGGCACAGGCGTTGCAGTTGCAATC
 CCAGGTCGAGGCGATGAGCAGGAAAAATCCGCAACAATGAAACCCCTCATCGAACAGTTCAAGTTGAAATCGCTCTGCTCA
 AACGCCACAAGTTTGGCAAGCGCAGCGAGCAATCAGTTTCGGCGCAAGGCGAGCTTGTGATGACCTGCTCGACACCGAC
 CTTGAAGCTATCGAGGCGGAGCTGAAACAACTCCTTCCAGCTTCGCCACAAGCCGAGCCACGGCAATCCCGAAACGTTT
 GCCATTGCCGCGCGAGTTCCCGCGCACGGTGATTGCCACGAACCTGAAATACCAATGCGCCTGCGGCTGCCAACTTC
 AACGCATCGCGGAAGACGTCAGCGAGAAGCTGGATTACAGCCGGGCGTGTTCACCTCGAGCAACATGTGAGGGGCAAA
 TGGGCTGCCGTCAGTGCAGAACCTGATCCAGGCGCGGTGCCAGCCAGGTTATTGATAAAGGCATCCCGACCGCAGG
 TTTGTTGGCCACGCTGATGGTGGCCAACTTTGCCGATCACTTGGCGCTGACAGACAGGAAAAATCTTTGGCCGCGCGG
 GGCTGCCAATGCCCGCTCGACCTGGCGCAGTGGGTGGGACAACTGGCGTGGGCTTCAGCCACTGGTCGATGCACTG
 CGTGAGCCGTGCTGAACAGGACGTCATCCAGCCGATGAAACACCGGTGCAATGCTTGCAACAGGCGAGAGAAAAAC
 CCACCGGGTCTATGTCTGGGCTACAGCACGACGCGTTTTTCGGCGCTCAAAGCGGTGGTTACGACTTCAGCCCAAGCC
 GTGCGGAGAACATGACGCAACTTCTAGGCGACTGGAATGGCAAGCTGGTCTGCGACGACTTCGCTGGATACAAGGCC
 GGTTTTGAAACAGGCATCACTGAAATCGGCTGCATGGCTCATGCTCGCGCAAGTTCTTCGACCTGCATGTCGTAACAA
 AAGCAACTGGCCGAACAGGCGCTGCACTCAATTGGCGGTTTGTACGAGGTTGAACGCCAGGCTCGGGACATGAGCAACG
 AAGACCGTTGGCGAATACGTCAGGAAATGGCGGTACCGATCAGCAAAACACTGCATGATGGATGTTGGCCAGCGCGAC
 CTGGTGGCCCAACGGCTGGCCACAGCTAAAGCCCTCGACTACAGCTGAAACGCTGGGGAGCGCTGACGCGCTACCTGGA
 CGATGGGGCTGTGCGCATCGACAACAATCAGGTGGAGAACAGATACGGGCGTGGGGCTCGGACGCTCGAAGTGGTTAT
 TTGCCGATCGCTGCGCAGTGGCAACGAGCAGCAGCTATCATGAGCCTGATCCAGTCCGCTCGCATGAACGGGCATGAT
 CCGTATGCGCTACCTGAAGGACGTGCTAACTCGCCTGCCGACGTTACGGTTCGAAAGACATCAGCCAGTTGCTGCCGCATCA
 GTGGGTACAGATCTAG

Protein sequence: (SEQ ID NO: 160)

MTSSPNLDQMTPEQLRALAAQALQLOQVEAMSRKIRNNETLIEQFKFEIALLKRHKFAKRSEQISSAQGSLLDDLLDLD
 LEAIEAELKQLLPASPQAEPRQSPKRSPLPPQFPRTVIRHEPENTQCACGCLQRIGEDVSEKLDYTPGVFTVEQHVGRK
 WACRCQETLIQAPVPAQVIDKGIPTAGLLAHVMVAKFADHLPLYRQEKIFGRAGLPIARSTLAQWVGQTVRLQPLVDAL
 REAVLNQDVIHADETPVQMLAPGEKKTHRVVWAYSTTFFSALKAVVYDFSPSRAGEHARNFLGDWNGKLVCDDFAGYKA
 GFEQGITTEICMAHARRKFFDLHVANKSQLAEQALHSIGGLYEVRQARDMSNEDRWIRQEMAVPI SKTLHDWMLAQRD
 LVPNGSATAKALDYSILKRWGALTRYLDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLIQSARMNGHD
 PYAYLKDVLTRLPLTRSKDISQLLPHQWVQI.

RL060

DNA sequence: (SEQ ID NO: 39)

ATGATCCGCATCGATGCGATCTGGCTAGCCACCGAACCGATGGACATGCGCGCGGCGACCGAGACGGCATTAGCCCGGGT
 AATTGCGGTGTTCCGTTGCGGCGAAGCCGCACTGCGCTTATCTGTTCCGCAATCGCCGGGCTAACCGAATGAAAGTGTGG
 TGCACGATGGCGTGGGCATCTGGCTTGCCGCGCGTGCAGTGAACCAAGCAAGTTCCACTGGCCCGGCATTGCCATGGC
 TGCAGGTCGAACTCGACAGCGAACAACCTCCAGGCTTGGTGCTGGGCTGCGGTGGCAGCGCGTGGCACAGGCGGTGT
 GATCAGCATGCTGTAA

Protein sequence: (SEQ ID NO: 161)

MIRIDAIWLATEPMDMRAGTETALARVIAVFGAAKPHCAYLFANRRANRMKVLVHDGVGIWLAARRLNQKFWPGIRHG
 CEVELDSEQLQALVLGLPWQVRVGTGGVISML.

RL061

DNA sequence: (SEQ ID NO: 40)

ATGCGCCAACGAAGCTCTTACCCGAAACCGTTCAAAGCCCGAGTTCGTTAGGAATGCCTGCAACCTGGGGCAACGGTGTG
 CAGTGTGCCATCAGCCACGGCATCAACGCCAATGTCATCCGCAATGGCTGACGCTTTATCGAGACCAGCCGTACCA
 CCTCGTTACAGCCTTTGTCCGCTGAAGGCCACCCCTAAACGGCCAGCCGAAACGTCAGTGTCTATTGAAGTGGCCATG
 GCCGGCAATGATCAGGTGAAATAG

Protein sequence: (SEQ ID NO: 162)

MRQRSSYPKPFKAQVVQECLEPGATVSSVAISHGINANVIRKWLTLYRDQVPASLPAPVPLKATPKRPAETSVLIELPM
 AQQMITVK.

FIGURE 30D

RL062

DNA sequence: (SEQ ID NO: 41)

ATGGCTTTATCTCTTATTCGTAAGTCTCACTGCGTCCGCTCACGAAACATCTCGGCGTTGAAACGCGATGCCAAACGCTT
GCAGAAGAACTCCTTTCTTGTGTTTGGAAACAGAAATATCCATCAAGGTTTGCCTAAATGCGGTAGCAGTTTCTCGCGGCT
TCCGCTCACTCGCTGATGTCGATAAACTGGAGCAGCACATTGGCATGAATAGAAGCGCTCCATTCTGGGTGATCCGTTGGC
CGCAACGATACACACAGGGGGTACTGGAAGCGCTATATTGTTAGACCTTGAATATACCGAGAATGGCCCCGTCGTTTT
TACTGGAAACCCAAAGCACTCTATACTTCCAGCCTTAGTCCTTTTCTTGAGCAAATGAGCTTTAAGAACTACCCGGAC
TAATCCTCATCGAAACAAAGAGACCTCAATCCAAACACCCATATATTGACGCAATAGAAAAATTAGAAGTCGAAGAA
ACTCTAAATAAATTTGATTCTTGAAGTTCGAGACCGAAACCTTCCCGTTTCTGCTTAGTACCGAGGCTCGTTGCTGGAT
CGAGTCAATTGTCAGTTTATGCCAAACGACATCCAAGAGGAAATACGTAATAAAGGATGGTCAACTCACTTAGAGATCA
GTGCATATGAGCATGCAAAGTCTCGTAATCAAGTATTTGGCTCCTCAACTTCCCTTGGCTCCCTTCTCTCCATAAAG
TCAGCGATCTATCAACTCATTTAGGCGCATACCCTCCCTTATGGATGACGCCATCCTCCTGCGGCAATATCTAAAGT
TGATATACGCGGACCTCCTCTCGAAAAAGCTCAGAGGAAACCTTACTTTATCTCATAAAAAAATTAGAGAATCGACAGT
TCCACACAGGCATTTTATGTGAGCATGAGAGTCGATGGCGGCGTATGTCGTAATCTTCCAGGAATGATCCGGCTAGC
GAGGTACTAGCAGGAGTTATACACTCGTACTTTTCTTGAAGCAAGTAGAGACCATCGCTCACCACCTTTATGTTTC
AGATGGAGCAGTTCCCTATGCTCCCAAGCTTCTAGGTTTAGGCGGCGCATCGGTCATTGCAATGGAATCACTGAAATTC
CCGACGGGGATGGTCTTGGGGAGTTCTATGGCTACAAGAACTCACTTAAAGTCAGCTCCTTATCTAACGAATACAGTTTC
ATGGGTAAGCATGTATCACTAAAGTAA

Protein sequence: (SEQ ID NO: 163)

MALSLIRSLTASASRNISALKRDAKRLQKNSFLVFGTEYPLKVCQNAVAVSRGFRSLADVLDKLEQHIHMNRSAFFWVIRG
RNDTHQVLEALYCLDLEYTENGPVFTGNPKHSILPALVLFLEQMSFKKLPLGLIETKETSITQTHIFDAIEKLEVEE
TLNKRFLDLRDLRNLPSVLSLEARCWIESI VSLLPNDIQEEIRNKGWSTHLEISAYEHAKSRNQVFGSSNFPVCFPLSIK
SAIYQLISGAYPLWMPSSSGEISKVDIRPPEKSSEETLLYLIKLENRQFHTGISCEHESRWRPYVVLFSRNDPAS
EVLAVIHSYFSWKQDRDHSPTLYVSDGAVPYAPKLLGLGGHTVIANGITEIPDGDGLGEFYGYKNSLKVSSLSNGIQF
MGKHVSLK.

RL063

DNA sequence: (SEQ ID NO: 42)

ATGAACGCTCTGACCAACCGCGCCCTCGCGCTCCACCTGAACATCAACCTGACCGACTTCATCGACGAGTTCGG
CGACGAGTCTCGGAGTCGCTCAATCGCTCCAACCCCCCGGTCTATACCGGCTCCGTCAACGCTCACCGCCAGTTGGTGA
TGGACCGACTCAAGCGCAAGCCCTTCGCGGCCAGGCGGAGTCTGTCAGGCCATCACCGCCCTGCTGCTGGACCGTAAC
GAGCAGGCCGGAATCATCAACGCGCGAGATGGGCACCGGAAACCATGATGGCCATCGCTGTGCGAGCGGTATGCAACGC
CGCGGCTATCGCGGACCTGGTCTCTCGCGCGCACTGGTCTACAAGTGGCGCGCGAGATCCTGGAGACCATCC
CAGCGCGCGCTGCTGGGTACTCAATGGCCAGATACTCTACTCAAGCTGCTCAAGCTGCGAGATCAGATGGGCGACGCC
TACGACGGGCGCCAGGAGTTCTTCTCCTCGCGCGGTGCGGATGCGGATGGGTTTCACTGCGCGCTCGCTGCTGGAA
GAAACGCGCGCGCGCGCAACTGCTCGCTGCGTGGCGGATGCGGACAGGTCTCGAGGACCTGGAAGGCAACCTGG
TCACGGTGGAGGAGTTGAGCGTGGTGACCGTCGACGTACCTGTCTCCTCGCGTGGGCGCTCTGGACGCTGATCCGA
CCAGGCAAGCCGACGCGCGCAACCGCGCGCAACGATTCTCAAGTCGATGTGCGCGATACCAACCATCGGCCCGGTGAG
GGCGGAGCGCTGCTGAACGACTTCGCGGAGGACTTCTGGCCACGATGTTGGTGGACAACGCTCTCGGAGTTTCATCAACC
TGATGGACGCGCAAGGGCAACTTCGTCTTACGCGATCGGCGAGGCAACGCGATGGAGCGATCGATGGCAACATCGAGTTC
GGCTTCGTTGAAGGCGGCTACCAACCGACCGAGTTCATCAAGCGCTACCTACCTGATGGCTACTTCGACCTGCTGGTGT
GGACGAGGACATGAGTACAAGAACAGCGGCTCGGCCAGGCGCAGGCCATGGGCGTTCTCGCAGCAAGGCACGGAAAA
CCGTGCTGCTGACCGGAACGCTCATGGCGGCTACGCCGACGATCTGTTCTATCTCCTGTTCCGATCCTCAACCGAGCG
ATGATCGAGGACGGCTATCGGCCCAACGCGCGCGGAGCATGGCTCCCGAGCCATGTCGTTCTATGCGCGACCGGTGT
GCTCAAGGATATCTACACCGAGCGGACCGGTGATTCGCAACAAGACAGCGCGGGGCAAGAAGCTCTCGGTACGACCGGTGA
AGGCTCCCGGCTTCGGCCCAAGGGCATCCACCGCTTCGATTGCGGTTACCGGTGTTCTGAAGCTCAAGGATATTGGT
GGCAACGATGCTGCCGACTACCAAGGAGGAGTTTCATCGAGTGGCCATGGCGCTGAGCAGGCTTCGGCTATCAGCGCT
GGCGGCCACGCTGACAGCGGAGCTCCGCCAGGCTTCGGCGGAGAGATACCAAGCTCCTGGCGTGGTCTCAACGTGC
TGCTGGCTTGGCGGAGTCTGTTTCCGACCGAGATCGTCAAGCATCCGCGAACCGGGACACACTGGCCTTCGTGCCA
GCGATCTTCGGTGACGAGCAGTTGATACCAAGGAGCAGGTGCTGGTGGACCTCTGCTTCGAGGAGAAAGCGAAGGGCCG
CAAGGTTCTGGCATAACCGCTTACAGCGGACGCGGACACACCGTCCAGGCTGAAGAAAGTCTCGAGCAATCCGGGC
TGAAGGTGGCAGTGCTACGTGCTTCGGTCGATACCGCTCGACGCGAGGATTGGATCCTCGACGAGTTCGATCGCGGCATC
GATGTGCTGATACCAACCGGAGTGGTGAAGACCGGGCTGGACTTGTCTGACTTCCCGACCATCGCGTTCTGCAAAAC
GGGCTACAACGTGTACCCCTGCAGCAGGCGCGCGGCGTCTGGCGGATCGGGCAGAAAGCACCAGGTGAGGCTGGTGT
TCTTCGGCTACGCGGAGCTCGCAGATCACTGCTTACAGCTGATGGCCAAGAGATCGCTGTGGCTCAGAGCAGCTCG
GGAGACGTTCCCGAGTCAGGTCTGACTCGTTGAACAGGATGGGGATTCTGTGGAGATGGCGTTGGCACGAACTCAT
CGCAGCATGA

FIGURE 30E

Protein sequence: (SEQ ID NO: 164)

MNALTQPAALAASHNLNLTDFIDEFGDELLESNRSNPPVYTGSVNAHRQLVMDRLKRPFAAQAEVVQAITALLDRN
EQAGIINAEMGTGKTMMAIAVAAMVHAAGYRRTLVVSPHLYVKWRREILETIPAAVRVWVNLNGPDTLLKLLKLRDQMGDA
YDGRQEFFILGRVVRMRMGFWRLACWKKRAAGGQLLAACPDGQVLEDEGLNLTVEEFERGDRRRTCSSCRGALWTLIR
PGKPDGNNRRATILKSMCRIPTIGPVRAERLLNDFGEDFLATMLVDNVSEFINLMDAKGNFVFSRQAKRMERSMANIEF
GFGEGGYQPTFEIKRYLPDGYFDLLVLDEGHEYKNSGSAQGQAMGVLAARKARTVLLTGTLMGGYADDLFYLLFRILTQR
MIEDGYRPNARGSMAPAAMSFMRDHGVLKDIYTERDGDSSHKTARGKKLSVRTVKAPGFGPKGIHRFVLPFTVFLKLDIG
GNVLPDYQEEFIDVPMAPQASAYQRLAATLTAE LRQALARRDITLLGVVLNVLLAWPDCCFRPEIVKHPRTDRTLAFVP
AIFGDEQLIPKEQVLVDLCFEEKAKGRKVLAYTVYSGTRDTSRLKKVLEQSGLKVAVLRASVDTARREDWILDQVDRGI
DVLITNPVLKVTGLDLLDFTIAFLQTGVNVYTLQQAARRSWRIGQKHPVRVVFYAGSSQITCLQLMAKKIAVAQSTS
GDVPESGLDSLNDGDSVEMALARQLIAA

RL064

DNA sequence: (SEQ ID NO: 43)

ATGGCCCTCATGTTCCCGCGCTTGGCGCGCAACTTTCACGCAACGGCTACTTCCCTACCGATGAGGTC
ACCCCTGAACGCGCTCTGCAGGCCCTCACTCTGCCCCGTCGGAAGGATGAGGATCTGTGACCCCTGCGCCGGTGAGGG
TGTTGCCCTGGCTGAGGCAGCACACCCCTCGGCCGCGATCAGGTCCAAGCCCTCGCTGTGAGTACGACCGCGAGCGCG
CCGACCATGCCCGAGGATTGCTTGACGAGTGTGACAGTACCTTTTCGACACCATGATCAGCAGGCAGTCGTTTCGGA
CTGCTCTGGCTCAACCCGCTTATGGCGACCTGGTGGCGGACCACTCCGGTGGCTGCGAGTACCAGGGCAGCGCCGCGG
GCGTCTGGAGAAAGCGTCTACCGAGCGCTGCCTGCCGTTGCTGAGTACGCGCGCGCTCATGGTTCTGATTGTTCTCACT
ACGCTTGGACGATGAGCTGACTGGCTGGTTGAGCAACCACTTACCGCGCTGCGCATCTACGACGCGCGGATCCTACC
TTCAAGCAGGTGGTGATCTTCGGCATCCGGGTCGTCGCGAGCACTGGCCCGGGCGGACGCCAATCAGGTGAGGTCTCG
CTGCAGCGGATCGGAGCGGGCCAGGAAAGGCCGAGGAAATTCAGCGGCTTGGCCGTGGGAACCTACGTGGTTCTGCG
CGGCCACCGAGCTGGAGCACTTCTACCGAGTAACCTGGAGCCGAGCAGTTCCGCGGTGAATCCAGCGGCTGCGA
GGTCTCTGGCCTGACTTCAACCTGCATCTCGCGCAAGCGGGCTGCGACCGCGCCACCAGTCCGCGAGCTGTCTCGCTG
GCACCTGGCCCTGGCCCTGGCCCGCGCGCATATCTGGCGTCTGCGATCGAAGTCGGGCGGATCCTGGTCTGTAAGG
GTGACACCTACAAGGACAGGTCCGCAAGACCGAATTACCGAGGACGACGACGCAACATCACGAGGTGAGGATCCTC
ACCGACCGTTTCATCCCGATCATCCGGCATGGGAAATGACACCTCCTCGGTCAATCAGGGCCGCGTGTGACCATCAG
CTCCTCGGCGCGACACGGAAGAGCTGAAGAGCCCCAACCTGAGCGGCGCCCGCACCCGACCGCTGTGATCAGCC
CTGGCCGGTCTGTAATGACCGCAGCCGTGAGCCACCTGGTGGAAACCGGTCAACTCAACCCAGCGCTTTGCTGAAACGC
CATCTGGCGGGAGATTGGGGAACGCTGGACAGGAAGACTGGAAACCAACAGAGAGCCCTGAAGTTCCGCGATCGGCT
GCTGTCTCTACGACATCGACCGCGGACGAATCCAGGCTCTGGATCATCACTGAGGCAGACCGCAGCTCAACCCACGC
TTTGTCTCCTAGCGATTACTGA

Protein sequence: (SEQ ID NO: 165)

MALMFPLRLARNFARNGYPTDEVTLERLALQALTAPSGRMRICDPGAGEGVALAEAAHTLGRDQVQALAVEYDRERADHA
RGLLDRVLHSDLFDTMISRQSGFLWLWNPYGDVLADHSGASQYQSGRRLEKAFYQRCPLLLQYGGVMVLIVPHYVLD
DELTGWLSNHFTGLRIYAAADPTFKQVIFGIRVRRQDLARADANQVRSRLQAIAGAQEKAEIIPAAWPPEPYVLPATS
ELEHFYRVLTLEPEQFAGEIQRLRGLWPDFNLHFAQAGLQPRFPVRELSRWHLALALAAGAISGVVRSKSGRILVVKGDY
KDKVRKTEFTEDDGNITEVRILTDRFIPIIRAWEMTPSSVNQGRVLTISSSAATTEAEPEQPEPAPAPAPLLISPGRV
VMTAAVSHLVETGQLNPAPLLKRHLAGDWGTLTDQEDWNTNQRALKFGRLLSSYDIDAGDESRLWIITEADRSSTTLLLP
SDY.

RL065

DNA sequence: (SEQ ID NO: 44)

ATGCCAGTCCACCCCGCTCTACCGATCGAAGAGTGTCCAGACCTGTACGTGACGCGCTGCGTGTGCGACGAGCAGTG
CAACCTGGTCTTTCTTTTCGGCCTGGGGCGCGACACCGTGACACAAGAGTTCCTGGCCAGGCTGACGCTGGGCGGGGAAG
AAAATGGCATCGACCATTTCCACATCATCGTGGACGGCGCGCTTACCTGTCTTCCAAACAGGATCTCCTGGAGAAA
CGCACACCCGTCAGTTCGCGCGACGTTGTTCCGACGCTGCTCAATCTTTGGCTGTTGATCGGCGCGCTCGGCGCC
CGACCGAGGCAATCACCTCGCCTTCGCACTCTGACGCGGATGAGGATCCACACAGAGGCTCTGGCCGCTGGTGATGG
AAACCTGTCCGCTCCCCCTCCTGACGACTGGCGGACCGGTGATGGAGGTTCTACCCAGCAGCAGATGTTGACGGCC
CTACCCGGGACGATCGGCAACGTCTGCGCCTGGCGACTCGCCCTCGGGTTCGACGTGCTTGAACCCACCTCGGTGAGGT
AATCCGCGAAAGCATTTTACACCGATGCTCAGGCGCAAGCCTGA

Protein sequence: (SEQ ID NO: 166)

MPSPPTPLYQIECPDLYVDACVDEQCNLVFLSAWGRDVTQEFRLRLTLGREENGIDHFHIIIVDGRRLPVFPNQDLLEK
RTTRQFRGTLFGSLLNLWLFDRRASAPDRGNHLAFALLQRDEDPHQLWPLVPMETCPLPLLQHWREFVMEVLTQHQLMTA
LPGTIGNVCAWRLALRVDVLEPTLGEVIREISILTDAQAQA

FIGURE 30F

RL066

DNA sequence: (SEQ ID NO: 45)

ATGAATCCATTGTTACCAACCTCACCAGGAAACCTCGCCTACCTCGAGGACCACTGTCCAACAACGACGTGCGCCG
 CGACGACGAGCTCATCGACTTGTTTCATCGAGGAGCTGTCGCTGACCTGGAGCAGGCGGAAGCGGCTGTGCGCTACGCG
 ATCAGTACCTCTGCCAGGTCTTCTGATCGGCCAAGGCGGCTGCACCAAGCCGATGGACTCAGCTTCGACCCTCACACC
 AAGAGCGTTCGGTAG

Protein sequence: (SEQ ID NO: 167)

MNPLFTNLQETLAYLEDQLSNNDVAGDDELIDLFIEELSLTLEQAEAAVALRDQYLCQVFLIGQGPHQADGLSFDPH
 KSVR

RL067

DNA sequence: (SEQ ID NO: 46)

ATGGGATGGCTTTTTCACATCAGACGAAGGACCTGCTGCGTGAGCTGCTGGCCCCAACAGTACCTTCGAGGCAG
 CACCGAGGTGCTGGCACACGAGTCTCCGGCAATGAACTTTGAGTGTGTAACGAACCTTTTCACCTTGGCGGATTCT
 ATTTCCGCAAGCCGCGGCTCACTCGATCACCATGATCGAGCTGCACCTGCTGGACTGCTCGGCCGGCAATGGGGCTAC
 AAGACCATTTCCGGAAGCGCGGCGCGTCTACTACGGCTGTCCGCTGGAGTTCCTGGACCTGGCTCAGATGAGATCAA
 CCAGGAATGGCGTAAACGCTGACGCACGAACACCAAGCCTGA

Protein sequence: (SEQ ID NO: 168)

MGWLFHQTKEDLLRELLAPTSTFAGSTEVLAHAVSGNELWTVVKRTFHLAGFYFGKPAGHSITMIELHLLDCSAGQWGY
 KTIPESAGPFYFGCPLEFLDLAHDEINQEWKRLTHEHQA

RL068

DNA sequence: (SEQ ID NO: 47)

ATGAAATCGATCTACAACACCCAGGCTTCAGCGAGGAGTTGTTGCTGGTTTGGCGCTCGCTGCGCGAGGTGGGACTGGA
 CAATCTGGCTGACCAAGTTCGCGCGGCAAGTGTTCGACCGATCCGTCGTCGACCAAGGCCATCATCGCACTGCGTGAGCGGG
 TGAAGACCCCTTCGCGGAGCATGCGGCCGACACGAGCCCTGGTTGTACTGCGACTGGCAGGCCAGGCAACAGCTTAC
 CGGCTCCTCCAGCGCCTTGAGCGCGCAACACGCTGA

Protein sequence: (SEQ ID NO: 169)

MKSIYNTPGFSEELLVLCASLREVGLDNLDQFRAAVFDRSVVDQAIIALRERVKTPSPEHAADNEPWLCDWQARQTAY
 RLLQLERATR

RL069

DNA sequence: (SEQ ID NO: 48)

ATCCCTCACCACGATCCCGCTTCGGCGGGATCATCCTTTTCGCGAGGTATACCATGATCACAGTTCCTCCGGACAGTTGGC
 CATTGGAACCATCAACGGTCGCTATGGCGAGTTCAATGTGGGAAACTCTGGACTTCGATCGGGGAGTTTCATCATCAAGG
 ATGCCCTTCTGGATCAACACACCGAAGGCAAGTACCGCGGTGATTTCGTATCGCCAATATCCGCCCCCACCCTACTCC
 GCCGGCGGTTCGCTAGTCTATCGAGATCCGCGCCATAGTGGACAGCATGACGCTGAACGATATGGACAGCCTCAGCGACGA
 GGAGGTAGAGCGTCTTTCCGGCAATGAGGTGGATCCGCTCGACGAAGTGCCTCGAGATCCAGCTCCCCACAGTAGTACCGG
 CGATACCAACAAAGTTCGCGCTCAGCCGAGAGTCAAGCTCTGTGCTCGCTGCAACCAAGGAGGACGCGCTTTTCGGTATG
 GACTCTCGGCTCCTCGAGAGCAGCCGCTCTCTGGACACAGACGCGGATGCAGAACTGTTTCGGACCGTCTGGCGCT
 AGGCGAAATCGTCAAGCTGGACACACCGTTCGACCGCAAGCGACTACGCCAACAGTGCCTGCGACTCGGCGCGCTGGGT
 ATGAGCTCGACTTCAAACAACAGGTGTGGACCCGCAAGGAGGCCGATGA

Protein sequence: (SEQ ID NO: 170)

IPSPRSRFGGIILFAGHTMITVPGQLAIRTINGRYGEFNVGKLWTSIGEFIIKDAFLDQHTGKYRGDFVIANIRPHYS
 AGGRLVIEIRAIVDSMTLNDMDLSDEEVERLSGNEVDPLDEVPEIQLPTVVPAIP

RL070

DNA sequence: (SEQ ID NO: 49)

ATGACCTCTCTCAACAACCACTCCAGCGCAGGTACACTGCTGCGTACCTCAAACCTCCGATCGTTCTCACCACGCGGC
 CTGGCTGCGCTGGTCTATCTCGCCAACCTGCCAGGGTCGACGAGATGGGCACCCGGCTGGCCAGTGTGCTTCAAACCG
 CCTGGCAGGAGCTTTCTCTCAGCCGACCGCAAGCACATCCAATTCCACCTGTACCAAGGAGGAAGAGGGGCGGAC
 CGCGCGCTCGCGCTGCTGGTCTCTCGATAGTCGAGCCGTCGATGAGCCTTCTACCTGCGCATCGAGTTGCAGGAAGA
 GTGCTTCGCCGAACCCCGTTACCGAGTAG

Protein sequence: (SEQ ID NO: 171)

MTSLNNHSSAGHTAAYLKLPIVLTAWLRLVYLANPARVDEMGRRLASVVQTAWQELSLQPTAKHIQFHLHYHKEEGQD
 RALALLVLSIVEPSDEPSYLRIELQEELAEHPVTE.
 PKSPSPQKSKPLCLAATRDAPFGM
 DTPAPAEQAASLDTDADAELFGTVWPLGEIVKLDTTVDKRLRQCVRLGALGYELDFKQQVWTRKEAA

FIGURE 30G

RL071

DNA sequence: (SEQ ID NO: 50)

ATGACTCAACTCAACCGTTTATTTCGCGGCTATGAGAGTTTCCGCATCGAGCGAAACCTGCAGATCACTGACGAAGGCAA
CAATCTACCGTCTACCGCGCTCTGCATGAAACCCAGCAGCACCTCCCAGACGAATATTTTCAGTGCGAGCTGTGCTACT
TCAATAACGATTTCGCGGTGGTAGTCCAAGAGTTAGACGATGAAAGAGTTGAAAAATGCCCTCACCAAGGAATAGTGAGA
AACGTACTTTACAGCATCTACGGTGAGCAGGACGGCAGAAAAAGCTTATCGGAGATCAATACTCACTGACCGAAGCCGA
GAGTGTCTGTTACGATACCTTTTCGTTTCGCGCGCGGTTATAACCCCTGCTGGGAGATCAGAAAAACATCTACCCATCAGCG
CGTGGAATAGCCTCTACGAAAGGTTCTCGACCAAGATGCCAATCCGCTTGCCTTCGGTGTGGTATCGCTCTTCTGGTGT
AACGAGCAGCGGTGGCGGCTTTCGCTTCGACAAACCCCTTGGACGGATGAGTGTCTGGAGATCCTGGAGATGACCGC
AGCCGCTCTTCGACAAAGACAGCTTGCCTTCGGCTCGACGAACACCTTGTGATCTGCTTCACCTCGCGGACAAAGCAG
ACATTCGCGCTCCTGGTACTTGATCCATTTCGCGCCACGCTCAAGGGCTGCGCTTTATGACGATTGA

Protein sequence (SEQ ID NO: 172)

MTQLNPFIRGYESFRIERNLQITDEGNNLPCYRALHETQOHLPEYFQCELCYFNNDFAVVVQELDDERVEKCPHQGIVR
NVLYSIYGEQDGRKKLIGDQYSLTEAESVVRYSFGGYNPCWEIRKTHLPISAWNSLYERFSTKMPIRLPSVLVSLFWC
NEHGAVGFRLLHNTPTWDECLEILEMTAAALRQEQALFGLDEHLVDLHLAQADIRLLVLDPPFAPTLKGLPLYDD.

RL072

DNA sequence: (SEQ ID NO: 51)

ATGGGACTGGTGTTCCTACCGAAAGGAGAATCACCATGCAATACGGAAAGCTGGCGCTCGCCCATCTCAGCCTGGAAC
GCCGTTCAGGTACTTATGAATAAGAACCGTGCTTACTACATCGGCACTTCTGACGAAGAAGGACCAGCCTCGCGCGAGT
CGGTTGAATATTACCCCTCAGCGCACTTGCCCAACAGGCATTAGACCACGGCACTTGGACGCAACTGGAATATTAA

Protein sequence: (SEQ ID NO: 173)

MGLVFPTEIRITMQYKGLALHLSLELPLQVLMNKNRAYYIGTSDEEGPASRESVEYYPSRELAQQALDHGTWTQLEY.

RL073

DNA sequence: (SEQ ID NO: 52)

ATGGGAAATGTTTGGCGATTATGCCAGGGCAGATACCTGGGCATTGTTGTTGGCCAGGAACAGCCAGGCGAAGTTGCAGA
ACTGACTGCTGAGCAGCAGCTCGTCTCGACGTCGCTGAGGCTAACCTCCTCAACTCCGGCAGGGCGGGCAGTTCTACG
ATTGGATGTTGCTCTATGATGATCTCCAGATAATGGAGAACACACGCCCTGGGGGGAGATGGTGGCCCCCGGATGGGTA
TGCGATGAAGAGTGGCGCATAGCGTAG

Protein sequence: (SEQ ID NO: 174)

MGNVWRLCQGRYLGIIVGQEQPGEVAELTAEQQLVLDVAEANLLNFRQGGQFYDLVAHDDLQIMENTTPWGEVMPGWW
CDEEWRIA.

RL074

DNA sequence: (SEQ ID NO: 53)

CTGACGGGCAAGGTGTTTCTCGCTTACGAACTGGAGAATCATCATGAGCAACAACACCCAAGCCCAAGAAGC
CAAGTATTTGACCTGCACACCACCGGTATCGGCTACCTCAATCGCATCCGCGAGGTACCGATCCGCGAGGTGAACCAT
TCCTCGCCGTAACCGTCGACGCCCTCCATGGCGCGGACAGCGTGAATACTCTACATCGACTGCAAGTGGTTCGGC
GCCCAGGCTGAAAAGCTTGTCCGCCGTTGCAAGGAAGCAGTCGAGGCCAAGAAGAAGGTTCTGATTTCCTTCCGTATCGG
CGATATCTGGGCGGATCCCTTCATCCACCAGAAAGGCGAGAAAACAGGCAAGCCCGACGCAAGCCTCAAAGGCCGGCTGC
TCTTCATCTCCTGGATCAAAGTGGATGGCACACCGTCTACGATGCGAAGGAAGAAGCTGAAAAAGCCAGCAAGGCCAA
GGCGAACCTCAAGGTGAGCCCGCAGCCCCGCTGAGCACGCTGAACAAGCCGCTGCTTGA

Protein sequence: (SEQ ID NO: 175)

LTGKVFLRFLRNWRIIMSNNTQAQEAQYFDLHTTGIGYLNRIREVPIRRGEPFLAVTVAALHGAADSVEYSYIDCKVVG
AQAEKLVRRCKEAVEAKKKVLSFRIGDIWADPFIHQKGEKQKPDASLKGRLLFISWIKVDGTTVYDAKEEAQQAQGGK
GEPQGEPAAPAEHAQAAA.

RL075

DNA sequence: (SEQ ID NO: 54)

ATGTCCAAGCAATCCACAGCTTCGAAATCGGCTTTGCCCTCGGCAGTGTGTGCGTGAGTTCGCGAGAGCGCTCAGTCG
CCCTCCGCTCGTAGTGCAAGCACAAGCGCGGTTGCGTTGAGAGTCCAGCGCATCGATCCTGCCTTCTTGGCCGGCCCGA
CCGCTGGCGAGCTAGAACACATCAGCGACATCCAGCCATCGTCCGGCTGAAGAAGGTCAACCTGAATGACTGGTATCTA
GCCAATACGCGCGAGGTGCAAAAGCCAAAGCGCGACGCAACCAAGCCGGCAAGGCGACCGCCAAAGCTGAAACGCC
AGTCAGGAAGGAGCTCAAGATGGGTTCCCTCGACCATTTGATTGCACCCAACTCCGAAAGCGAAATGGGAGGCCCCCTC
TCCAGTTAGAGTCCCTGAACGATCATGAGATTGCTCTTTGCCAGCACCTCCTGGTAGCGAGTCTCTTGGGAACCTCCAT
CGCGTACTCAGGAGCAATACCAACAACGCTGGCAGGACTACTTGTCCACCATGACGGATGAACAAGTAGCTGCTCTCGG
CCGCTAA

FIGURE 30H

Protein sequence: (SEQ ID NO: 176)

MSKQSTSF EIGFALGSVVREFRRALSRPPVVVQAQAPVALRVQRIDPAFLAGPTAGELEHISDIPAIVRLKKVNLNDWYL
ANTREVQKPKRKRPKAKATAKAETPVRKELKMGSLDHLIAPNSESEMGRPPLQLESINDHEIALLPAPPGSAVSWELH
RRTQEYQQRWQDYLSMTDEQVAALGR.

RL076

DNA sequence: (SEQ ID NO: 55)

ATGGTGTTTCTCCTGCAGGTTGAGGGTGCGGAGAAAACACTGGCCCTGGCGGGGAAGTGGATTCCCCGCTGGGTTGCGGA
AGGAGCTTCTATCGACCGAGGCCGACCGCGCTACAGAAAGCTATGCGGTCTTGGGTTGGATCAACACGGTGGGCT
GTGCTGCAGCATTTCCGATCCGAGCTGCATGGGGGCTGCTGCTGACAACGTCAGCAGATCACGCGTTCATCATCGAAGC
GGGGGGCGAAAGTGTCAGGTCAAGCAGGAGGGGGAGCGGATGCAGCGGGCGGAGAGCGAGGGCGGAAGAGCGCGGCTGG
TAGAAACCTGTCAAAGGTTTCCCCAGCCGTGTCTGGAAGGGGAGTCAAGTGAGCCACCTGTGGTTGAATCGTCGATCCC
TGGGCATTGATCGTCTCGATCCCATCACCGGCCATTTATCGTGGCTTGGCCAGCAAACAGTAGGCACGCGATCCGCGTACA
AAGGGAGCCCTGCGTATCACCGGGGGCCACCGGCAGGAGAGGATCCCGATGGGTAGCCTGATAGTCTGGAGCAGGA
GCATCAGGCTACCCATGGAGAGGGGAAAAGGAGGGGCCGTAACACCAGTACGACCCTTAAATCGAGGAAACACCGAACCT
CTTGA

Protein sequence: (SEQ ID NO: 177)

—MVFLQLQVEGAETKLALAGKWLPRWVAEGSFYRPRPTDRATRSYAVLGWINTVGCAAARIRAAWGHVADNVSRSRVHHR
GGRKCGQAGGGADAAGGERGRKSAAGRNPVKGFPSRVWKGSSQVSHLWLNRRSLGIDRLDPITRPLSWLQQQTVGTHPR
KGLRITGPPAGRRIPMGLIVLEQEHQATHGEGKRRGRNTSTTLKSRKHRTS.

RL077

DNA sequence: (SEQ ID NO: 56)

ATGCCGCTGATGTGGATCGTCTGGTGCTCGCGCTCATCACCGGACCTGGCTGAGTGTACAAAGCGACCACGCGACCTC
GAGCGCCGAACCTGCCCCAGGTCGACACCCTGGCCAGGAGCTTGCTGCTCTTCCGGTCCAGTCTGGCGGAGTACGCACAG
CCAACCCCGGTTTACCGGTTCCGCCGGGACTCCGCTCTTGGTTTACCGGCTGGTTCGCAAGCCAGCGCGGCTTCAG
GGCTACATCGCGCCCGCAGCTACGCTTTCATCGCTCGCGCGCGGGGCTGGCGCGCGCGCTGGATGCTGGTAC
GGAATCCGACCTGGTTGGCGTCAGGCGCAACGGCCAGTTAGTCACGCGCGCGCTCGGAGCCACTGTCATTGCGCTCCCTA
CGCCCATCCCCGAGGGCGCGGTGGTCCGCGTCAAATAA

Protein sequence: (SEQ ID NO: 178)

MPLMWIVLVLALITGTWLSVQSDHATSSAELAEDTLARSLLLFRSSSLAEYAHANPGFTGSPADSALGLPAWFRKPARLQ
GYIAAGTSYAFIASPPAGLAAAVDAGTESDLVGVRRNGQLVTRRLGATVIALPTPIPEGAVVAVK.

RL078

DNA sequence: (SEQ ID NO: 57)

ATGAGGAGTACGCGCAGCAGTGGATTCTATCTCGATCGAAGTATGATCGCCCTCGTCTGATCGCCATCGCGACCGCCGG
TGGCATATCGGTCTGATGAGCTACCTGGACGGCTTGGACGAGCAGCAGCGGCCAGCAGCAACAGCAGGTGGCCAAGG
CAGCGGAGAAGTACCTGAAGGACAACTTCAGCAGGTTCTGGCCAGCGCCGGCGCCACGGCCCCGGCGGTGATCACCGTC
CCGATGCTGCGCAACACCCGTACCTGCCGAGGCTTCCGCGACACCAACATCTACGGCCAGCAATACAGGTCTTGGC
CCGCAAGCCGGCGGCCAACAGCTCGAAACGCTGATCGTGACCAGGGTGGACAGGTAGCTTCCGAACCTCTCGATCCGCC
GGATCGCGCAGCTCATGGGAGCCACCGGGGCTATATCTCGAAAACCAACACAGTATCGCCAGGGCGCGCGCTGGCAG
GTGGCTTAAAGCAATTTCCGTAGCGCTCCCGCGCTGGACATCTGGCGACGGCGCTGTTCTTCCAGGACGGCGCCATCGC
CAACGAGTACCTCTACCGCAATGCCGTCCCGGGTATCCTGAACTCAACCGGATGAATACCAACGCTGGACATGGGAGGCA
ACAATATCGCCGCGAGCCGGGGGATCACGGCCAGCGGCAACATCACACCAGCGCGGACATCAGCGCGCGCAACGTGACA
GCCACTGGTACGGTGAAGACCGGCACTGCTGACGTCGCCGGCGAGACGTACACCGGAGGCTGGTTCAGGACCCGTTGGTGA
CACGGGCTGGTACAACGAGAAATGGGGCGCGGCTGGTACATGAGCGACAGCACCTGGGTGCGCTCTGGATGAACAAGA
ACGCTTACACCGCGCGGAGATGAAAGCGGCAAACTCACCGCCGAGGGCCGGACGGAAGTCGGCGAGTACCTACAGCTC
AAAGGCGTGGCCACCGAAGGAGCCAACTGCTCGCCGAACGGGCTGGCAGGCATCACACGACCGGACTCTGGCTGTCCTG
CCAAAACGGGAAATGGGGACGAACCGCGCTCCATGCGCTGAACACACCGCGCGGCTGATCAAGGACTGGTGTACGT
TGCATGGTCAGGATAGCGCATGGTGAATACGACTACGCTCCGCTACGCGATCACTGCGGCGCGGATTCTGCGCAGTG
GGCTTCAACCAGACATTGGCACCACCTACTCGTTCGGGCTAATCACTGAGATCGGCCAGGCTTCAACTACCGGAACC
CTACAAGACCCCCGACTCGACCAACGTGACCGTTACCTGCGTGAAGTAG

Protein sequence: (SEQ ID NO: 179)

MRSTRSSGFISIELMIALVVIAIATAGGISVLMYSYLDGLDEQHAAQQQQVAKAAEKYLKDNFSTVLASAGATAPAVITV
PMLRNTRYLPAGFRDNTNIGQQYQVLARKPAANQLETIVTTGGQVASELSIRRIAQLMGATGGYISKNTNSIAQGAAWQ
VALSNFSGAPGAGHLATLFFQDGAIANEYLYRNAVPGHPELNRMNTTLDMGNNIAAGAITASGNITTSADIARNVT
ATGTVKAGTADVAGETTYGGWFRTRGDTGWYNEKGGGWYMSDSTWVRSWMNKNVYTGEMKAGKLTAEGRTEVGEYQLQ
KGVATEGANCSPNLAGITSTGLWLSCQNGKWGRTAASMRNLTAGVIKDWCTLHGQDSAMVNYDYVRYAITCGGRFCV
GFNQTFGTNYSFGLITEIGPGFNYPEPYKTPDSTNVTVTCVN.

FIGURE 30I

RL079

DNA sequence: (SEQ ID NO: 58)

GTGAGTGTGAACCCGATCATCCAGGCTCAGTTCGTGACCTCTACCTCGGTGAAGGCTTCGCCGACGTGAAAGGCCTGGC
CGGCGCCGGCGCGCCGAGTCGAAGTGCCTCGCGAGTGGGAGTCGCACGTCCAGGAAGTGTCCAGATCTGCAGGCAAA
CGCTGGAGGAGCTGCAGGACCTGAGTTCGCCATCGCTCGACGGCGTTCTGCTTCGCGTCACCTCTCTCGAAGACGCT
TTCAGTGGCAGCGTCTTCGTGCTGCGCCGGTCGAGCGCCCAATTGCGGGAGTTCAGAGATCGGCTATCCGAGCGAAGT
GGTTTCCGCACTGATGGATCCGAGTTGCAGGGCTGGTCTGTCTGCGCGGAGATGGCGACAGGCAAGACCAGCTCCG
CCGCTCTCTGCTCTGGCCCGCTGCAGGAGTTGGGCGGGTGGGCTGCGCGCTCGAGGACCCGAGGAAACCAACCTC
AGCGGTGAGCATGGGCTCGGCGCTGCATCCAGGTGAGAACCTCAGGCGCTCAGGCGGATACAGCGAGGCCCTGCTGCG
CACGCTGCGGGCGCGCCGACCTGGTGTGATTGGCGAGATCCGCGACGAGGACACCGCTACCAGGCTGCAAGGCTT
CTCTGACCGGACGCTGGTGTGTCGCCACCATTCACGCGAAAAGCTGTATCAGGCGATCGAGCGCTTGGTGACGCTCGCC
CAGCACTGGCGAGAAACGCTACGACGTGGTGGCGAAGGCATCCAAGCTGTGATCTGCCAAGCGCTGGAGAGCGATGG
TTCTCTCGCGCCGCTGACCGCGAGCCAGCTGTCTGCTGCGGACGACGCGCCGCTCCATGCGCGACAAGATCCGCCGAA
AGGAGGCTCATCTGCTGCAGGACGACCAAGCTCGCCAGTCCCGCAAAGCCTATGGAGATAA

Protein sequence: (SEQ ID NO: 180)

VSVNPIIQAFVDLYLGBGFADVKLAGAGARRVEVPREWESHVQELLQICRQTLEELQDPEFAIVVDGVLRLVLTLEDA
FSGSVFVLRSSAQLREFQEI GYPSEVVSALMDPQLQGLVLFCEMATGKTSASLLARLQELGGVGCavedPQETNL
SGQHGLGRCIQVRTSRRSGGYSEALLRLRAGADLVLI GEIRDEDTAYACKASLTGSLVIATIHAKSCHQAIERLVTLA
QPLARNAYDVVAEGIQAVICQALES DSGSSRLTAEPLLF TGDGSPMRDKIRKKEAHL LQDDQARQSRQSLWR.

RL080

DNA sequence: (SEQ ID NO: 59)

ATGAGCACTACGCAACGCACTTCCCGTCCGACGCGAGGGCGGTTTCGTTTCCATCGAGATGATCATCGTGTGATCATCAT
CGCCATCGGGGTGCGGCTGGGCTGGCCGACGCGCTGGAATGTTCAGTTCTCCAACGCCAACGAGGAACAACGCAACA
TCAGCGTCATTGCGGCCAACGCGCGCCCTGAAGACCTCTTCGGGCTACGGCTCCAGCGGTACCAACCTGATCCCCAGC
CTGATCGCAATCAACGGCGTGCAGGAAGCATGAGTGTCTCTCCGCGCTCGTCTACAACGCTCTACGGCGGATCGGTAC
TGTCTCGTCCACCGGCATGGGCTTCTCGATCACCACAGCAAGTTGCCCCAGGACGCGCTGTATCAGCTGGCCACCAAGA
TCGGAAGAACACCTTCGAGCAGACCAAAATCAACAGCGGATCTCTGATCACCGGAGAAGTGACCACCGCAGCCGCGACC
CAGGCTGCAGCAGCAGCAACAGCATTACCTGGACCTATAGTTCGTGA

Protein sequence: (SEQ ID NO: 181)

MSTTQRTSRPTQGGFVSIEMIIVLIIIAIGVGLGLAAAAGMFSSSNANEEQRNISVIAANARALKTSSGYGSSGNTLIPS
LIAINGVPKNMSVSSGVVYNVYGGSVTVSSTGMGFSITTSKLPQDACITLATKIAKNTFEQTKINS GSSITGEVTTAAAT
QACSSDSNSITWTYSS

RL081

DNA sequence: (SEQ ID NO: 60)

ATGGGGGCTTCTGGGAGCAGTTGCAGTTCGCTTCTACAGCAAGCAGTTTCGGCCGCAAGGAACGCCTGCAGTTCTACGA
AAGCATGTCCACCTGTCTGAAAACGGGGTCCCGTTGAAGGATGCTGTGGCAGAGGTGCACAAGATCTTCGCTCATGAGG
GGCAGCATCCGTTTCATCCGGTGGCCATCGCCAGTCGCGAAGCGCTGATGGGGCTGTCCAACGGCAAGCGTCTGGCCACC
GCCATGGCGCTCTATCTCCCCGCCAGGAGCGAGCGTTGATCGAGGCCGCGAGATGAGCGGCAACCTGGTTCAGGCCAT
GGGCGATGCGCTCTCCCTGGTCGAGGCCAGGCCAGGATCCGCGCCACCATCTGGCAGGCGCTGCTTACCCTCGGCGC
TGTCGCCCATGATGGTGTCTCTGCTGTGCATCGTGGCTATCGCATGGTCCCCAGCCTGGCCAGGCTCTCCGACCCAGTC
ACCTGGACCGGCCGCTCGCCAGCTCAACGCCATTGCCAGCTTCTGTCACAGGACCTGGTATCTACGTTCTGGTTCGCGCT
CATCACCTCACGGTGGTGGTTCATCGTCACGTTGCCGACCTACCGCTGGAAAGGCCGGGTCTGGCTGGACCGGACGCTGC
CGCCTGGTCCATCTACCGCATGCTCCAGGGCACCACCTTCTGTGTAACATGGCGGTCTGCTCAACGCCGGCATACGC
CCCTACGACAGCTGGCCAGCATGATCAAGATCTCCCCGCCCTGGCTGAAGCAGCGCTTGAAGCTGCCGCTACGGCGT
GGGCTGGGCCAGAACTTGGGTGTTGCCCTTCGAGCGCGGTCAGATTTCCTCCGACCGCGAGGCCATCCAGTACCTGT
GCATCTCGCCAAACCGGGAGGCTTCTCCGAGCGCTGGTCAAGTTCAGCCGCGCTGGCAGGAGACCAGCTCAAGCAG
ATCGAGCTGGCGCGCGGCTGGTGAAGAATTCGCCCTGATCTTCATCGGCGCGCTGATGATCTGCTGCTCTCGGCGC
CTACCAGGCACAGCAGCTCATCAATCCATGAACCTGA

Protein sequence: (SEQ ID NO: 182)

MGGFWEQLQFAFYSKQGRKERLQFYESMSTLLENGVPLKDAVAEVHKIFAHEGQHPFHPVAIASREALMGLSNGKRLAT
AMALYLPQERALEEAGEMSGNLVQAMGDAVSLVEAQARIRATIWAQALYPSALSAMMVFLLCIVAYRMVPSLARLSDPV
TWTGPLATLNIAISFVTGPGIYVLVAVITLVVVIIVTLPTYRWKGRVWLDRTLPWWSIYRMLQGTTFLLNMAVMLNAGIR
PYDSLASMIIKIPSPWLKQRLAARYGVGLGQNLGVALRSAGHDFPDRQAIQYLCILANRGGFSEALVKFSRRWQETS LKQ
IELAAGLVKNFALIFIGALMILVLLGAYQAQQLIQSMNH.

FIGURE 30J

RL082

DNA sequence: (SEQ ID NO: 61)

ATGACGAACCTTCAGATTGCCGCGCTTGCGCAGCCCTCCATGGTGACCCAACTGCTCACCGCCGACGGTGGTGAATGGGA
 GGTATCGAAGCACCTGCAGGAAATCATGGCTCTGGCTGCCGACGGCAGCTCTATCTATCGGAGAGCCACCAGAACGACA
 TACACGTTCTGTCTGTTTCATCGACCGTCTCGATCGCCGCTGGCTTCCGATACCAGCTCAACCTCACCGACCTGCAGACCAT
 CACCAGCTTTACCGCGCCGCTCGCCATGGACGGCTGGTCGATAGCGATGGCCAGCGCGCCACCCAGATGCAGGAGCGCGT
 GGTCAAGATCATTGTAAGGCCACTGAGCTGCGCGCCAGTGACGTGCATTTTCGTCTGAGTCCCGCCGCGACCGGCAGCA
 AGATCCGTTTCCGCGCTCGACGGCTGCTGAAGACCGTGCAGCAGTTCCGCGACCGAGGAGCTGCACGAACTCTGTGCAACC
 ATCTACCAATCCATGTGCGACGTGGCCGAGCCACTGTTCAAGCCGCAACTGGACAGGACGCGCGGATGAGCCAGACCTT
 CGTCGAGAAGCTCAACCTGTTCACTGCGCGGATCGCCACCCGCGCTGCGGGGGTTCCTGATGATCCTGCGACTGC
 TCTACGACGACACCGGCTCGACAGCCTGGAGCAGCTCGGTACCTGCCCGAGCAGAACGCACTGTTTCGATCGCATGATG
 CGTATGCCCTACCGCATCAACATCTGTGCGGCCCCACGGGTGAGGAACTCGATGACCTTGAAGGTACCCCTGGAAGG
 CCTCGACAAGCTCCATGGCGGATCCAAGCACATCCTGACCATCGAGGATCCGCGGAATACCGCATTCGCGGCGAAGGCA
 TCAACAGACCCCACTGCTTACGACGCGACCGACCCAGACGCGAGAACGCCAGGCTGGGCCGCGGGCATCGCCAACGGC
 ATGCGCCTGGATCCGACTACATGATGATCGCGAAGTACGCGACCTTTCGCGCGCTGTCGCGCGTTCGTGGTGCATG
 GACCGGCGACCGGCTATGCTGACCTGACACCAACAGCGCATCGGCATTGTCCAGCGCCTGAAGGACCTGGGCGTGC
 ACCCGGCTGTGCTGTTTCGATCCGCGCTGCTGACCGGCTGATCAACAGAGCCTGCTGCCAAGCTGCCCCACTGC
 AAGTTCGCTTCCAAGAGCAACCAAGCAACTCGCGCCGACTTGGTCGAACGGGTCCGACGCTTGACCGATGTTTCCCA
 GGTTCAGTCAAGGGGCTGGCTGCGAGGCTGCGGTGGCTCCGGGTCAACGCGCGCTCGATCGTCGCGAGGTGGTTC
 TGCCACCTCGCTTCAATGCGTGTGTTTCGCCAAGGCGCGCCAGCCGAGGACGCAACTACTGGGTCAAGACCATGCGAG
 GGCATACCAAGCAGCCACGCCATCCGCGCATCAACGAGGGCATGTTTCGACCGCAGATGGTTCGAGGATTTCATTGG
 GCCACTCGACTTCGATGAGCATCTGCTCGACGACAGCTTCTACTCGCAGGAGGCGTGTGA:

Protein sequence: (SEQ ID NO: 183)

MTNLQIAALAPQSMVTQLLTADGGWEVSKHLQEIIMAAADGTLYLSESHQNDIHVLSFIDRLDRRGFRYQLNLTDLQTI
 HQLYRAVAMDGLVDSGQRATQMQERVVKIIRKATELRASDVHFVSPAGTGSKIRFRVDGLLKTVEQFRSQELHELCA
 IQSMCDVAEPLFKPQLDQDARMSQTFVEKLNLFSAIATPRAGGFLMILRLLYDDTGLDLEQLGYLPEQNALFDRMM
 RMPYGINILSGPTSGSKSMTLKVTLLEGLDKLHGGSKHILTIEDPPEYRIRGEGINQTPLVYDATDPDAERQAWAAGIANG
 MRLDPDYMIGEVRLDFAVAARFAMTGHGLWSTLHTNSAIGIVQRLKDLGVDPLLLFDPAALLTGLINQSLPLKCPHC
 KVRFDHQQDLAPDLVERVRLTLDVSVHVHKGPGCQACRGSGVNGRSIVAEVVLPTLAFMRVFAKGGPAEARNYVWKTMO
 GITKHAHAIRINEGMFDPMVEDFIGPLDFDEHLLDSDFSYQEQAC.

RL083

DNA sequence: (SEQ ID NO: 62)

ATGCGAACTGAGCCGATCGGCATGGCGTGGCCGTGCTCTCTCTCTCGCGTCTGGCCAGGCTGCGCTGGCACCCTGG
 CGAACTTGCGGAGATCCAGGCCAGGCCATCTCACCGAGGCCAAGGTGCGCTGGCCACGGCGCAGCGGCAATTGGGAAG
 GCAAAGGCGAAACCGGCCAGGTGCTCAGCGCCAGGGGCGAGCTTTCGCGATGCCGCTGCGCGCGCGCCGCGACGATC
 ACGCAGCCGCTTCCGCCAGTGGTGGGACCATCTACGGCGCGCGCGCAAGATGACTGCCACGTTCTTGTTCGCGGCGG
 GTACGAGGTTGACCGCGCAGCGCGCGGAGCTGCTGGCAAAATACCGCGTCGAGTCAATCTCGCTGGACAGGTCGTGC
 TCAACGACAAGGACGCGCAACCGCGTGGCGTGGGCTTCTCCAGCGTTGCCCCACCAAGCCTCTACGCGCCCAAGG
 GCCTCGGTTCCGCGCGCGTGGCGGGGTGTACCGCAGCGCTTCATTAGTAG

Protein sequence: (SEQ ID NO: 184)

MRTEPIGMAVAVLFLLASGQACAGTVGELAEIQAQAILTEAKVRLATAQRQLEGKGETGQVVSAGGQTFAMPVPAAPPTI
 TQPVPPVVRTIYGAGGKMTATFLFPGGYEVDAAAGAEPLGKYRVESI SLQVVLTDKGNRVPVGVFSSVAPTQASSTAQG
 ASVPPALPGA VPPFIQ.

RL084

DNA sequence: (SEQ ID NO: 63)

ATGGAAGACCTGACCTCGGCAGCGTGGACAGACGCTCTCGATCCTGAGCTACCACGGCAACAAGTTCGTAGCGCGCT
 GTTCTGGCGCCGCTGTCCAGCCAGCGGCGAGTACATGAAGGAAGCGCGAAGCTGGGCAAGGAAGAGCATCTGGACATCG
 TTGCCATCCGCCATTACCCAGCGGTGATCCAGGCCGCTTCTGTTTCGAAGTCGAAGGCGCAGTCAAGGGGATGTACTCC
 CTGGCCTCGCGCTTTTCAAGCCAGTTTCAGCGGCACTTCTGGCTGCTGGAAGTTCGACGAGGACCGCTACGCGCTGGT
 CGCCACGCTCGATGGCGGATTGTCCCGGGCAGGATCTGGTACACCCCTCGACGAGGCCCGGACCGGTCAGGAAGC
 TCTCTACGCGCGCGTGTGCGAAACGCAAGGTCTTCTGTTCCGAAGGTTTCGATTTCCTCGTCAAGGACTTCGACATC
 GAGGAAGTGTGCGCGCAAGCGCTGCGGCGCGACTACCGCTCCGGCAACTCACCTTCGCTTGTCCGCGAGGAGTG
 GACGCGAGTGGCCCTGCTCGGTTGCGTGGTGGTGGCTAACCGCTACTACCTATGGAATGCCACCCAGGAAGAGC
 TCGCCAGGCAAGCGCGCTCTCGAGGAGCAGAGCGCTCGCGGAGCTGGCCGAGAGAACCGCCAGGCAAGCAGCGC
 CTGGACCTGGCGTATTGCGAAGCCTTGGACGCTATACCTGACCTCGAGGACATGCTACGCGCTGTAGCAAGGCAAC
 GGGGTACTGTGCTGTGATCCAGGGCTGGCTCTTCAATCCAGCAAGTGCAGCGGAGGTCCTGGTTCGCCACCTACC
 ACCGTACCGGCAACAGCAGCAGCGCCGACTGACAGCGGCGGCTGACAGCGGCTGTTTCGCGGACCGCCCCGCTTCGTCATC
 GACAACGGCAACACCGCGGCTGAAAGTCTGATCTGAAGGTTGCCATCGGCGAGTGTAGGCGCTACTGCGCGGAGCA
 CGTTCTGACGGCGCTGACGAGCCACTGTACCGTCAAGGGGTGAGCCCAAGCTGTGATCAGCCAGGAGACAACCTCCGC
 CCTCCCTGGCGGGAAGTGCAGTGAACAGCAAGTGGTGTTCCTTCTGGAAGAAATACCTTCAGCGCCAGACC
 CGGCTCCCGGAGACCTGACCTTCCAGGGGTGCGCGTGTCCGATCACCAACCTCGAAACCGCTCAAGGA
 CAGCAGTGGACTGGACTGTACAGGAGAACTATGCGAAGTGA

FIGURE 30K

Protein sequence: (SEQ ID NO: 185)

MEKPDLGSRGPDVSLSYHGNKFVSGLFWRPLSSQRYMKEARLKGEEHLDIVAIRHSPTVIQAGFVSKSQGAVKGMYS
LASALSGQFDGDFLACWKVDEDRYALVATLDGAIVPGQDLVTTLDEARDVRKLSTRGVLRNAQVFVPEGFDFFVKDFDI
EELLAPKRLRRDYRLQLTFGLSAREWTAVALLGCVVGGSLTAYYLWNAHQEELARQAALLEEQRRRLAEAEKNAQAKQP
LDLASLQKPWTLPDLEDMLRACSKATGVLSLSIQWLFESSKCDGRVLVATYHRTGNSTAADLTAAASQHLFADRPAFVI
DNGNTAALKVDLKVAGSDEPLLPADDVLQALTSHLYRQGVPEPKLSISQETTPPLPGAEEATEQQVVLPSWKKFTFSAQT
RLPADLTFQGLPAAGVRITNLETTLKDSQLDWTVTGEIYAN.

RL085

DNA sequence: (SEQ ID NO: 64)

ATCGTGTGCGAAGCTACGGCAGATTCCGCGTCTACGATCGCAGCGCAGGTGCGCAACACCCGACCGGATCGGCGCGATAC
GGTGGTGTCTCCGACAAACCTGGGTGAGCAGCAAAACCCCTAAGCGTTTCGCACACCTTGTCAGTGACTGCATCGTGA
CGTGGCGCCCTGCAGGCGCAGCGTCGTGTCAGGAGGCCCGCCAGGAAGTCATCAACCAATGCCACATGGCGGTGAGTATC
ACGCCCGCAGCGCTGAACCCGGCGCCCTTCGCGTGCAACCTCAGCAGCGCGCGAGCAACGCCCCGCGCCCATCCAAGG
CGGCCAGGACATGGCCACCATGCTGTTCTCTGCCTCCGTGCGCAACCGCATGTCGCTCGGTGCGCGCGCGAGCATGGGGT
CGAGCTTCGGGTCTACGGTTCGCGGTCTCTGTACAACATCAATGGAACGGCAAGTCAGCGGGTCTCTCGATCTCATC
GCCGCCCGAGCCGCGGTCTCTGGCGCTACAACCCAAACGAGAAAAGGGTCGAGTTCTACTACCTGGACACTCGGACCTT
CCGCGATGTACGCTTCGACGACGTCAACACGGTGGACTCCACCGTGCGTTCCGGTATGACGACGGCCGCGCGCATCAGCG
GGGACGGCTCCGGATCCACCGGACAGAATGGCAGCTCCGGCATCAGCGCGGACTCCGGCAGCAAGCAGACACCAGCTCG
GAGCTGAAGACATCGATCCTCAGCGACATCGAGAACAGCATCAACTCGATGCTGACGCCGAGCATGGGACGCATGTCGCT
GTCGCGTGCCACGGGCACCTGACCGTCACCGACCGTCCAGAACTCCTCAACCGTGTCCAGCAGTTGGTCAACCGAGAGA
ACGAGAGCATCACCAAGCAGGTGCTGCTGAACGTCAACGTGCTCTCGGTGCGCCTGACCGACAAGGATCAACTGGGGATC
GACTGGAACCTGGTCTACAAGTCGCTCAACAACAAGTGGGGCATCGGCCTGAAGAACACCATGCCGGGCATCGATCAAAG
CGCGATCTCCGGCTCCGTGAGCATCCTGGATACCGCCAACAGCGCCTGGGCAGGATCCAAGGCCATGGTCCAGGCGCTGG
CCCAGCAGGGCCGCTCTCGACCGTCCGATCCCCGTCCGTGACCAAGCTCAACCTCCAGTCGGCGCCGATCCAGATCGGC
CGCTACGACAGCTACCTGGCCTCCAGCCAGATCTCCAACGTGCGCCAGGTGCGCAGTACCACCTCGCTGATCCCGGGCGC
CGTGACCAAGCGGCTACAACATGAGCCTGCTGCCGTTCTGTATGGAAAGCGGCGAGATGCTGCTGAAGATCAACATCAACA
TGACCTCCCGGCCGACGTTGAAATGCAGACCAAGCGGGGACTCCAAGCCAGTTCCCGAGCTACGACATACAACCTGTTTC
GACCAGAAGGTACGTCTGCGCAGCGGCGAGACCTTGGTACTCTCCGGCTTCGACCAGACCAAGGACACCAACAAGGT
CGGCACCGCGCAGCTGGGTCTTTCGGTCTTGGCGCGGGCTGACCCGCAATACCAAGCGCAGGTGTCATCGTGGTGTCTGA
TCACCCCGCTGCTGCTGGGCTGA

Protein sequence: (SEQ ID NO: 186)

IVCEATADSASTIAAQVRNTRPDRRTVVVSDKPWVSTKPLSVSHTLSSDCIVTWRPAGAASLQEAQAQEVINQCHMAVSI
TPDALNPAFAVQPPQRRASNAPPPIQGGQDMATMLFPASVANGMSLGGGSMGSSFGSYGPRSLYNIKWNGKVSGLDLI
AARAGVSWRYNPTEKRVFYLDTRTFMYAFDDVNTVDSTVRSGMTTAAGISGDGSGSTGQNGSSGISGDSGSKQTTSS
ELKTSILSDIENSINSMITPSMGRMSLSRATGTLTVDTRPEVLNVRVQQLVNRNENISITKQVLLNVNVLVSVALTDKDLGI
DWNLYKSLNNKWIGLKNTPGIDQSAISGSVSIILDTANSWAGSKAMVQALAQQGRVSTVRSVPTTLNLQSAPIQIG
RYDSYLASSQISNVAQVGSTSLIPGAVTSGYNMSLLPFVMESEMLLKININMTSRPTFEMQTSKDSKAQFSPYDIQLF
DQKVRLRSGETLVLSGFDQTTEDTNKVGTGDAGFFGLGGGLTRNTKREVIVVLITPVVLG.

FIGURE 30L

RL086

DNA sequence: (SEQ ID NO: 65)

ATGACCAGGCAGTTGACCACTCTCACGCTGTGCTGCTGCTGCCAGCTGCACGACCCACAAGGCTGAGCCGGCCAGGCC
AGCCTTCGACAGCAGCCGCAATCCAGACCTGCTTTCTCCGACCTGTATCCAAACGGTGTGCAGCCGGAGAAAGAGCCCG
TAGTGGCTATGGGCGCTACACCTGGTCAGCACCAGCCTGATGCCGCTCAACGCGACCTGATGGCCAGATCATCGAC
GTAACCATCCGTCGAGCATGAACCCGAGCGTCAAGGACGCCATGCAGTACGTGATGAGCCGCTCGGGTTACTCGCTGTG
CCCGGCAGACGCCGCTCATGTGAACATCCTCTACACCCGGCCGCTGCCGCGAGCTCAGTACAAGCTCGGCCGATGACCC
TGCGCAACACCCCTCCAGGCTCCTCTCCGCGCCAGCCTGGCAGGTTAAGGTGCAGAGGTGCGCGCGGAGGTTCTGCTTCGTG
CTGCGCCCGGCTATCAACTTCCCCCGCGCGAGGCCGAAACCGGTCCAGCACTGTATGCGAAGCCCGCTGCCCAAC
TCCGCGCGCGGTAGCGCAACCCCTCTCCACGGAGAAAGTCAGCAGCTGGAGTGCGCCATCGTGGTGCCTCGGTGCCGA
CACCGCGCGGATCACACCCAGCCACGCTCCGCGCAAGAAGCTGAATCCACCACTGTGCTCCCCCAGCCGACCCGCGC
AAGGATGGCCACCCCTCTCTCTCCGCGCTTCGGCACCGACCAAGCCTGCGGCTCCGCCGTGAAGTCCACGCCCGC
CACTCCACCCACCGTGGCTTCGCGCCACCGGTCAAGGTGCTCAGCGCGCGGAACCGAGCCGCGCTGGCACAGGCCT
GGTCAGCCGAGACGGGATCAACCTGCGCGACACCTTGGAGCTTGGGCAAAGCGCGACGCTGGACCGTCCGCTGGGAG
CCGCGAGTCTCAACTATCCGATCGAGGCTCACTGACCTTCCAGGCTCCTTCGAGGACGCGGTATCCGAGCTGTTCC
CCTGTATGACGCTGCCGAACGCCCTTCTCGGTGAACGCGAGCCGCGCAGTCCCTGATCATCATCAAGGAGCGCAAGA
ACTGA

Protein sequence: (SEQ ID NO: 187)

MTRQLTTLTLCLLASCTTHKAEPARPAFDSSRNPDLLSPDLYPNGVQPEKEPVVRYGRYTLVSTQPDAGQDLMAQIID
VTIPSSMNPVKDAMQYVMSRSGYSLCPADAGHVNILYTRPLPAAQYKLGPMTLRNTLQVLGSPAWQVKVDEVARQVCFV
LRPGYQLPPAPRPKPVQQLYAKPAAPTTPPAVAQPSSTEKVTLESPIVVASVPTPAPITTSAPAKKPESTTVLPAAAPA
KDGHPSSPPAASAPTSPAASAVKSTPPTPPTVASAPPVKVLTTPPEPSRPLAQWASAEAGTSLRDLTLEAWAKRARWTVRWE
PQDLNYPTEAPLTFHGSFEDAVSELPFLYDAAERPFLVNASRPSLIKERKN.

RL087

DNA sequence: (SEQ ID NO: 66)

TTGAGCTTTAAATACTATTGGGCTAAATTTTCTGGGAGCTTTCTTCTTTGTTTTAGTCGCTTGAAAGGCTCCGTATT
TCCAAGCCTGGCATCAGTTAACCCCTTGGTAGTGGCTGGATTCACTACTATCCTGTTCTCTTCTCGGTAAGGCTTGTG
AAGACTTCGCTTTAAATATACGGAAGAAAGAGTTCTGGGTACAGGTTTTTCTCCGAAACCCCTGCAAAACAGGATTG
TATGCAGTCTTTTATTGGCTTGTATTGTTTTCAATTCCCTTGGGGATGATTTTTTTATTCTATAAATACGGAAGGC
CTCGTAG

Protein sequence: (SEQ ID NO: 188)

LSFKYYWAKFFWGAFFVLVANKGSVFPPLASVNPVAVAGFSTILFPFSVRLVEDFALKYTEKEFWVTGFFSETPAKTGL
YAVFYLACYLFSIPLGMIFLFYKYGKAS.

RL088

DNA sequence: (SEQ ID NO: 67)

ATGTCCAATGACAACGAAGTACCTGGTTCCATGGTTATTGTGCGACAAGGTCCAGACGATCAATACGCATACGAGGTTCC
CCCTATCGATAGCGCGCCGTTGCCGGGAATATGTTGGCGACTTGATTCAAAGAGACATATATCTACAGAAAAACATTT
ATTATCCAGTCCGATCCATTGTTGAACAAGGAACAAAAGAAAAGAGGAGATCAACAAGAAAGTATCTGATCAAGTCGAT
GGCTTGCTAAAGCAGATCACTCAAGGAAAAGGGAGGCCACAAAGGCAAGAGCGAGTCGATGTCATGTCCGGCAGTCTTGCA
CAAGATGGAATCTGATCTTGAAGGATACAAAAGACCTTTACCAAAGGCCATTCAATTGACTACGAAAAGCAGTCAAGCC
TCTCCATCTATGAGGCTGGGTCAAGATCTGGGAGAAGAACTCTTGGGAAGAAAGAAAGTACCTTTTTCAGCAGCTT
GTTAGAGATGAAGTGAAGCGGCGGTTGCTTACTACAAACAAGATTCACTCTCTGAAGCGGTAAAAGTGCTAAGACAGGA
GCTCAACAAGCAAAAAGCGCTAAAGGAAAAGAGGACCTCTCTCAACTGGAGCGGGACTACAAAACAGAAAGGCGAATC
TCGAGATGAAAGTACAATCCGAGCTTGATCAAGCGGAAGTGCTTGCCTCCATTGGTCAGTCCAACGCCAGAGCAATGG
CTTGAACGTGCCACAAGACTGGTTACGCAAGCAATTGCTGATAAAAAGCAGTGCAGACCACAAACAATACTCTTATCAA
GAATGCCCAACCCCTTAGAAAAGCAGAAAGCCATCTACAATGGTGAGCTACTTGTGGATGAGATAGCCAGTCTACAGA
CCCGCTTAGATAAGCTGAACGCCGAAACGACACGACGAGGACAGAAAGCAGAAAGCGGCGGAGGAACAAGCGTTG
CAAGATGCTGTTAAATTTACTGCCGACTTTTATAAGGAAGTAACTGAGAAATTTGGCGCACGAACATCAGAGATGGCGCA
CCAACCTGGCCGAAGGCCAGGGGAAAAATATCAGGAGTTCCGCGGAAGCAATCAATTCTGTTGAAAAACACAAGGATG
CGTTAAATAAAAACTTAGCCTTAAAGATAGGCAAGCCATTGCCAAAGCCTTTGATTCTCTAGACAAGCAGATGATGGCG
AAGAGCCTTGAAGAAATTTAGCAAAGGCTTTGGAGTTGAGCAAGCTATTGACGCCGCCAGCCTGTACCAAGAGTTCAA
GATATCTACGGAACCCGGGACTGGAAACCACTTCTTGTAAAAGTTGAAACACTAGCTGCTGGTGCAGGCCGCCAGTTGGC
TTGTGGGTATTGCATTGGCCACGGCAACGGCACTCTATAGGCATCTTGGGGTTCGCACTGGTAATGGCAGTTACCGGG
CGGATGATTGACGAAGCCTTCTAGAAAAAGCAAAACCTTGTAAATGTCCATTAA

FIGURE 30M

Protein sequence: (SEQ ID NO:189)

MSNDNEVPGSMVIVAQGPDDQYAYEVPPIDSAAVAGNMFGDLIQRDIIYLQKNIIYPVRSIVEQGTKEKKEINKKVSDDQVD
GLLKQITQKREATRQERVDVMSAVLHKMESDLEGYKKTFTKGPFFIDYEKQSSLSIYEAWVKIWEKNSWEERKKYPFQQL
VRDELERAVAYYKQDSLSEAVKVLRLQELNKKALKEKEDLSQLERDYKTRKANLEMKVQSELDQAGSALPPLVSPTEQW
LERATRLVTQAIADKKQLQTTNNTLIKNAPTPLEKQKAIYNGELLVDEIASLQTRLDKLNAEETRRRTEAERKAAEEQAL
QDAVKFTADFYKEVTEKFGARTSEMAHQLAEGARGKNIRSSAEAINSFEKHKDALNKKLSLKDRQAIKAFDSDLKQMMMA
KSLEKFSKGFVVGKAI DAASLYQEFKISTETGDWKPFVVKVETLAAGAAASWLVGIAFATATATPIGILGFALVMVAVTG
AMIDEGLEKANNNLVMSI.

RL089

DNA sequence: SEQ ID NO: 68

ATGAACCGTCCACGCCTGGTTAATCGTACCTCCGCGACACCTTCGACGCTTCTGCAGCGGGCTATCTTCGACGGCTACGA
CTTCGGCTTGAAGATCCCTACATCGCAGGCAGCAATCGCGCGCTGCTGGAGCTGTCCGGCTTCTTCATCAGCGCCCGG
AGCATCCGTTGCACCGCTACTGGCGGGTCCCAAAGGCAAGCTGCTGCTGAAGTGGACACTCTGTACAAACCGTCTCGCC
GAGCTAGCTGGAGGCCTTCACTCCAGTCTGGCGGGAGTTAGCTCCTTGGTCAATCCGCGCAGGCCTCGCTTGACCG
ACAGGCCTTCACTGGGGATGCTGCTGCGCATCGCGCCCTGGCCGAGGGCGGGCTCTACTGTACGGCGAGTTCCATC
CTGGTGTGTGGCGGTGGCTCGGCGGATGCGCGGGTATTCTGCGCCCATCGAGTCTCTGGCGCATCGACACCACTCCC
GAGTGTCTCCGAAGCAACCTGATTCTGGAGCTTGGCCTCGCCGAGGAACAATCGAGATTCTGGATACTGTCCAGGAGCT
GCTCAGCGACGGCAGCTTCGCGCCGTCGACCGAGCTGCCAGCATGAGCATCGGCGGTCCACAGCAGGAACCGGCAGCGC
CATCCCTGGAGGACGATCAGCCTCTGACATCTACCTCGCCGCGGTGCGGAGATCGAGCGCACCAGTACAGCTCGGCT
GATATCGAGGCGGCGCTTCAAGGCTACTCTACTAGGCCACCAAGCTGACGGCATCGCTCATCTGCTGCAGAGAACCAG
CGCCTTATTGGCGCAGACATGGGATTTGGGCAAGACCCGCGCAGGCGGTTCATCGCCGCTTCGATCCGCGCGGGCGGAG
CAATCCTGGTTCATCACCTGGCTACCTGCTGATCAATTGGCAGCGGAGATCCAGGAGGTCTATCCCTCGGCCACCGTG
GCCATCCAGCAGGACACCCAGAGGCGCAGTGGATCCTCGTCAACTACGAGCAGTTGAGCCCTTCGTGCGCAACGCTTC
GCGCTTCGCGGTGATGGTCATCGACGAGGCGCAGCGGATGAAGGAACGACGCGCAATGCACGCGGCACGGTTTCGACA
TTGCCGCCCAAGTGCCGAACCGCTACCTGCTTACCGGCGCAGCGGTGCTCAACCGCGAGACAGAGCTGCACACCTGCTG
CGCCTCTCAGGCCACCCCATCGGCCAAGTGGCGCTGAAGAGTTCTGCGACCGTTTCGCGGCAACCCGAGTTCCGCCA
GAGTCTGCGGGCGGAGCTGGGTGACTGGATGCTGCGCAGGCGCAAGATGTGCTGCCAGCCTCAAGGGCAAGCAGCGGC
AGTTGCTGAAGGTGGCCCTCTCCACCGAGGAACGCCAGCAATACGACGTGCTGCGCCTCGAGGACCGACCGGTCTTCGCG
CGACTCGGCGCGTTCGCGGCTTACCTGGAACCGGTGAAAGTTTCGCGTGGCGATGGACCTGTGAGCGAGCTCGACGAGA
GGACAAGGTGATCCTGTTCTGCGAGTTCAAGCCGACCGTGGCTGCGCTGAAGGAACCTGCGAGCAGGCGGACACGGCT
GCGTCACGCTGGTGGGAATGACTCGCTCACCAAGCGGCGAGAGGCGATAGATCGCTTCCAGCAGGATCCCGACTGCCGA
GTGTTTCATCTGCACTACGGCGGCGCAGGGACGGGCAACACCTCACTGCGGCGAATACGTGTTTTCTCGGCGTGGC
CTGGACTCCCGGTTCAGCAGGAACAAGCCGAAGACCGCGGTACCGAAACGGCCAGCTCCGCATGGTCTGGTGAAATCC
CACTGGTCGAGGCCACGATCGACGAGCAACTGTGGCAACTGCTCAACGCGAAACGCGAGGTTGCCAGGACCTCATCGAG
CCCGAGCAGGTCGACGGAAACCGCGCGCTTTAGCCGCAAGCCTAACTGGATAA

Protein sequence: (SEQ ID NO: 190)

MNRPLVNRTSATPSTLLQRAIFDGYDFGLKIPYIAGSNRALLELSGFFISAREHPLHRYWRVPKGKLLPELDTLYNRLA
ELAGGLHSQSWREFSSLVESAQASLDRQAFWGMILLRIAPLAEGGVLLSGEFHPGVVAVARRMRGVFLRPSSSWRIDTTP
ELLRSNLILEGLAEQFEIILDVQELLSGDSFAPSTELPSMSIIGGPQEPAPPSLEDESASDIYLAAPVEIERTEYSSA
DIEAALQGYSLLAHQPDGIAHLLQRTSALLADDMLGKTRQAVIAASIRAAGRPIVLITLATLLINWQREIQEVYPSATV
AIQODTPEAQWILVNYEQLSPFVANASRFVAVMVIDEARMKEPTAQCTRHGFDAIAQVFNRYLLTGTPLVNLRETELHTLL
RLSGHPIGQLPLKEFCDFRAGNPEFRQSLRAELGDWMLRRRKDVLPSLKGKQRLKVALSTEERQQYDVLRLIEDRPVFA
RLGALRRYLETVKVRVAMDLLSELDAEDKVLFCFEPKPTVAALKELCEQAGHCVTLVGNDLSLTKRQKAIIDRFQQDPDCR
VFICTAAAGTGNNLTAAANYVFFLGLPWTGQEQEAEDRAYRNGQLRMVVVKIPLVEATIDEQLWQLLNKRQVAQDLIE
PEQVDGNRALLAASLTG.

FIGURE 30N

RL090

DNA sequence: (SEQ ID NO: 69)

GTGGCACCTCTCGACAACGCCCCCTAGCGGGCCGCTACAGGATCCATCCCTGGCCCGCTACAGCGAGCGGCAGCTCGC
 CGTCGCCAACACCTGGGCAACACATTTCTCCTCGCAGGGACAGCTCGAACCAAGTTTCATCCGTCACTACCTGCGCAGCA
 CATCTACGACCAGGTGCTGGTGCATCACAGTCGCTGCTGACAACGGTGTGCGTTACACCATCATGCGTGCAGGGCCGCTA
 CTCCAGGTATTCGACGGTCAACTAATTGGTGCCTGGGAGTGCAAGCCTGCCATCGTATCCCGGCAAGCACGCCGTCTCG
 AGCAGGGGCTTGAAGCTGCTACAGCGCCTTCAAAAGTTTCGACGACGAGTGTGCTGTACTCAGCTCATACAAAGCGAG
 CGCAGACCTAGCCACAGATGGCCAGGGACGATCTCGGACTTCAACATCGCCTCGTGTATCCGAGCCACAGCAACAG
 CGCTACTACGCGCCAAAGGACCACTTCTACTTGAAGCAGATCGGAGCGGTCTTGCAGACCTTACAGACAGGTCTGGACCA
 AGACCTGCTGTTCCGCATCCGCTCGGTTCCGTCCTCTCGCCCCAGCTCTACAACCTGGCTGGCTCAAGGCGACCAAGTGC
 CGCGGCTGCAATGCTGAAGGCTCAGCCGGTCTTACGCGGCTACTGGTGGATTGCGAGGAGGAGTCTGGCCTCACACG
 ACGACCAACGACAAGCGCGAGAGCATCCGCCATTACCTTCTTGGCCCTTCCCCAGCTTGACAGTGAACGACCGCAGGC
 CGCCGCATGCCATGCGACTTGTACCTCGATATGGGCGTATTCTTGGGAGGTGCGGACGAAGGAATTTGCGTCATCA
 ACTTTTTCGCTGGCTATTTTCAGCGCGCGCGGCGCTCGATTGATTTCTTAGTCACGTCACTCCCGGCGGTGCGGGAGGA
 GCTCTCTTCCATCGCAACGGGAAGGCGGACATTGGGATGGCATGCTCTCTACTGGCGGCATCGCTAGGTAAACGGCG
 GCGATCACTCGCGCTCAATGGACAGCATTTCTATGCCGCTTACAATGCGATCCCTTGGCAAGTTTCAACACGCAAGCCG
 ACTACAACCGTCTCTTCAACGGCTGCCGTCGGATTGGCAGGATCCGGCATGGCTTGCAATCACTGACCGGCTGAGAGAC
 ATCAAGGAGTTCTATACGCCCTCGACAGGGGAACCTCACAGGTTGTTTCGGCAGGCGCGCAGCGCCCTGAAAGCGTATCT
 GGGTCATTGTACCTACCGACAAGCTGGCAACCTGGTGGACGACTACCACAGGTCCAGAGGAGCTGCGTGCCGCACTG
 AGAGCAGCCTGCCCCATCTGGTGCACACGACGAGTACACACCTGGGAGGGAATGCTGTCTGTCGGTCTTATCGATTGC
 CCTAATGGACTGCGATCGTCGAGCTCCGCTGCTCCTGCCACCTATATGCCGAACATATCGCTCTGGCACATTCGATCGA
 TAGCTACGACAGGCGCCCTACCGAGGAGACTGCCGACTGCTCTCAGTACGTGAGGCTGGTCTGCTCGCTGGCCTCTGCCG
 AATTGGAGCTCAGGCGTGAGCATGGCGAGCCTATAGGTAGGCGCTGGAGTCCCAAGCACTTTCCACGGTGCAACTGCGC
 GAATTGCGATAATGCCCCGTGCCGACCGACTCGCCTGCCGCCAGGCATACCGCTGGTTCATGGAACGAATTGCTCTGG
 AGCCATAGCGACGAACCTGAACCTGGCCGACATGACCGTCCACATGACGCGCTTCGCCAATGGTTCGCTGGAAGGCGGGCC
 TCGCCGAAGCCACGGCGAAGTGGTGTCTCACTCAGTTGGAAGACCGATGA

Protein sequence: (SEQ ID NO: 191)

VAPLDNAPPSGGLQDPSLARYSERQLAVANTWATHFSLAGTARTKFI RHYLRSTSTRWCITVAADNGVRYTIMRAGPL
 LQVFDGQLIGAWCKPAHRI PASTPSRAGALKLLQRLQKFD DAVAVLSSYTKRAHDLATQMARDDLGLQHLVYP SHSNK
 RYYAPRHQFY LKQIGAVLRTFRQVLDQDLLFAIRSVRCLSPQLYNWLAQGDQVRRLLQMLKAQPVLT PLLVDCEEGVPH
 TTNDNGESIRHYLPCFPQLDSERPQAAAMPQCDLYLDMGRILQVAD EGISVINFFAWLQAPRASIRFLSHVSPGRAGG
 ALFHRKREGRHSGWHALLAASLGNRRPITRAQWTFAYAYNAI PWQVHNAKPDYNRLFNGCPSDWQDPAWLAITARLRD
 IKEYTALDQGN SQVVRQARSALKAYLGHCTYRQAGNLVDDYHQVQRELRAAVQSSLPDLVDTDEYTTWEGMLSVGLIDC
 PNGLQIVELRC PADLYAEHIALAHCIDSYDQAA YRGDCRLLSVREAGRPLASAELELRREHGEPIGRPWSPKHLSTVQLR
 EFDNAPVPTDSPAGQAYRWFMERIRSGAIATNLNWPDMTVHMT RFANGRWKAGLAETAKWLLTQLEDR.

RL091

DNA sequence: (SEQ ID NO: 70)

ATGCGAAAAGAGAATATATCTGCCGAAATCACAGAGCGAGCTTTTGATTTTTCTATTGGTTCTCGCGATTGAGTTGAG
 CCTCAAAGAGAATGGCTACTTAAAAAATTACAAACCTGGAGCTAGGGCAGAGCCGGGATGGGAAAATTTGTACAAAACC
 ATTCTGACAAATCTCTCTTCCCAATCAGCCACAGCACTAATCGAGCAGAGTCCAGAGCAACAAATAGTCTGCCCCGT
 AGAGAGCTGGGTTGGCGTCCGGTTAAATTAGATGAGGACAAAAGCGACTTAGCTAGAGTCGCTCGCTTACTTAAGACCGT
 GCGAAACAATCTATTTCAGGAGGCAAGCATGGTGGTCCCACTGGGACAACCCAGCGAGGACAATACATCTTATTTCTT
 TAAGTAAAGCTATCCTTGACGAGTTTGTGCACTAGGAGACTTTGAGGCTGACTACAAGAGAATTTACTGA

Protein sequence: (SEQ ID NO: 192)

MRKENISAEITERAFDFYWF SRFEFSKENG YLKNYKPGARAE PGWENFVQNHSDKYSLSQSATALIEQSPEQQIVLP
 RELGWRPVKLEDEKSDLARVARLLKTVRNNLFHGGKHGGANWDN PARTIHLILLSKAILDEFAALGDFEADYKRIY.

FIGURE 30 O

RL092

DNA sequence: (SEQ ID NO: 71)

ATGCACATCGTAATCATTGAAGCCCCGGGCAAGCTGAAAAAGCTGAGGTCCCTTCTCCCTCGATTGTCCTCCGACGTGAC
CTGGCAGGTGAGGGGACAGCCGGCCACATCAGAGACCTACCGTTACGGGGCAGGATCGCAGATGCTCACCGTCGGCG
TGGGCCAGGATTTCAAACCGCACTACCAGATCTCTCGGGCAAGGAAAAAACCGTCGCACGGCTGAAGGAGCTGCGCAG
AAGCCGTGGAATCTACGTCGCATCGGACCCGGATCGCGAAGGCGAAAGCATTGGCTGGCACATCTCCAAGCTGCCGG
GATCAAGAACTACAAGCGCGTTGCTTCAAAGAAATCACAAGTCATGCATCACCGCCGAACCTCAGCTCGCCGCGTCGCC
TGGACCTCCCGAAGGTCGCTCGCAGGAATGCGTCCGTCATCGATCGCTGGTGGGTATCTGGTACCGCAGCGTTG
CGGCGCGTGATGGGTAGGCCGACCCCGCGCGCGTGTCAGTCCGTCGCGGTGTACCTGGTGGTCTGCGAGAGCGGGA
GATCCGCGCCTTACAGCAATCAAGCACTTCGGGGTGAAGTACCTTCTGTTCCGCCAGCGACGGCCGTACCTGGACGG
CGGAATGGGATCCAGTGCCCGTGTGTCAGCGAGGAGTTCCCGTATGTCCAGGATCGTCAACTCGCAGAAGTGGTGGGG
GCTATATCGTAATGTCATCGTCGAGACCTGCATTGATAGCGAAGAAACCGATGCGCTCCGGCACCGTTATCTCTCCCTC
GCTCCAGATGGCGCGCGGAATGCGCTGAAGTGGTCACCCGACAAGACGATGAAGGTGCGCCAGCGGCTGTATGAACAGG
GGCTCATCACCTACCAACCGGACGGACAACCCCAATATCTCGAAGGACTCGATGCCGGATATCCGTGCTGTCGCAAGGCC
TTGGGGCTGAAGTGTGTTGAGCAACAGCGGATGTTCAAAGCGGACCAAGACGCCAGGAGGCCACCCCGCCATCACCCC
TCCGACTGGATGGTCCGCGCGGTGAAACTGCTGATGACGAGCGCTGTACAGCTCATTGAGTCCGCGCGCTTG
CCAGTCAGATCGAAGTCCCGTGTACGCACTGAGAACCATCACCTCCTGGGCGTCGGCCCCGACAAAAGCCGCTGCGC
TTCGGCGCAAAGGGGAGCTGTTGAACGTCGCTGGTGGAGAAAACCTGCTGACGGGTGATGACGCCGAGGAGCAGAAGAA
CGAAACGCTTCAAACCCCATCCGATCCCGCGCTGGAGCCACGCCAGATACTCAAGGTCTACAGCGCGAGGTCTTG
AGAAGAAAACACCCCTCCCAAGCGATTACCCGACGCCAGCTGCTGGGGCAGATGAAGCGCCGCGGGATTGGTCGGCCA
TCCTCTACGCTCGATCGTGAAGAATCATCGACAAGGGCCAGGTGCAGATGAAGGGGCGAAGCCTGATCCCCGGCGA
GCTGGGAGAGGCCACCATCGCGCTCTGGAGCACAATTCAGCTTCTCAGCTCGACTTACCCGCAACCTCGAGGTGCG
CCTTGGACCGGATCGCAACAGCGAGGACACCTACATGAACGTGGTCCAGCAGTTCTACCAGCTACTACAGACAGAGCTG
CAGACACTCCGCGCGCTCCCGAGCGCACAGGACGAACACCGCGCAAGCTCCACCGCCAGTATCTCTCGCGCGCCGACAG
CGACTTCTTTGCGGCAAGTGGCTGCTGCCCCGTTACCGCAAGAAAGCCGCGCAAGGCGGCTTCGACTTCTGGGGTT
GCAGCGCTATCGAAACAGGGTGCAAGGTTAGCTACCCACCAAGAGCGCGCGCTGACTTCGACAACCCGCGCGGG
CTATAG

Protein sequence: (SEQ ID NO: 193)

MHIVIEAPGKLKLLPSIRPDVTWQVEATAGHIRDLPHVHGQDPQMLTVGVGQDFKPHYQILSGKEKTVARLKLRLQ
KAVEIYVASDPRREGESIGWHILQAAGIKNYKRVAFKEITKSCITAEISSPRLDLPKVASQECRRVIDRLVGYLVTPEL
RRVMGRPTTAGRVQSVAVYLVVLREREIRAFATIKHFGVELTFVSPSDGRTWAEWDVPVPVFASEFPYVQDRQLAELVG
AIRNVIETCIDSEETDAPPAPFISSSLQMAAGNALKWSPDKTMKVAQRLYEQGLITYHRTDNPNISKDSMPDIRAVAKA
LGLKCEVQQRMFKAQDAQEGHPAITPTDWMAAAAGETADEQALYQLIRVRALASQIEAAVYAVRTITLLGVGPDKKPLR
FGAKGKLLNVPGWRKLLQGGDAEEQKNETPSNPIPIPALEPRQILKVYSGEVLEKKTTPPKRFTDASLVGEMKRRGIGRP
SSYASIVKNIIDKGQVQMKGRSLIPGELGEATIALLEHNFSLSLDFTRNLEVALDRIANSEDTYMNVVQQFYQLLQTEL
QTLRALPSAQDEPRASSTASISSAPTSDFLCGKCGPLPVHRRKAGKGGDFWGCSGYRTTGCKVSYPTKSGRPDPDNPRG
L.

RL093

DNA sequence: SEQ ID NO: 72

ATGGATCAAAGCCTTTGCACATGCATGCCAACGCCAATCGTCAACCCCAAGGAGCTGCGACTGTGCCACATGTTAGTCGG
TAGAATTTCCCGATAACATTGATCGCAGCGACCATTTGGTTGAGCTATGACGGCAGCGCTGGTGGGTGATGCGGATG
AGCCCGCGACGGAGGACGAGGTGGCGGCTCTGTTGGTCAAGGCTGGTGGTGTCACTACGTGCTGGTGGGATAG

Protein sequence: (SEQ ID NO: 194)

MDQSLCTCMPTPIVNPKEKRLCHMLVGRFTFPTLIAGDHWLSYDGSAAWVDADEPATEDEVAALLVKAGGVTTWCWG.

RL094

DNA sequence: SEQ ID NO: 73

GTGGCAAGGGCTTCCGAATCGGAATCTCGACCAGTACGAGGTGCAGTGTGTCAAAGAGAGCGACCGATACCGACAAGCT
GGACAGACGACACTTCAACGATCCCAACCGGACTGTACGGGCTATTGGTGTGAGGCGCGCGGAAAGGGCTACGGGTGT
TCGACTGCCCTACAGTCATCTCGCATGCGGGCGTCTGGTTGAAAGGGTTTGGCCAGGACGAGCAACAGCAGCTCGAC
TTCTGA

Protein sequence: (SEQ ID NO: 195)

VARASESEISTSRCSVSKRATDCLKDRRHFNDFHRTVRAIGAEAAARKGLRVFDCPYSHPMRASWLKGFAGQEQQLD
F

FIGURE 30P

RL095

DNA sequence: (SEQ ID NO: 74)

ATGGCTACCCCGTCTTCTGGGAAGCCAACATTGGCTCGGCGCCGAGCACCAGCTTCCCCAACGGCAACAATCCCC
GCGGCAGTTGCTGCGACTGAACGTGATGTTTCGACAACTCGATTCCCCGATGGCCAAGGTGGCTACAAGGATCGCGCGGCT
TCTGGTGACGCGTGAATGGTGGCATCAGGATGCCAGCGCTTCGCCGAAGTGTTCAGAAAGGTATGCGCGTCAAGGTC
GAAGGCAGGGCCATTATGGACCGCTGGCCGGACAAAGAGTCAGGCGAAGAAGTCCAGGCGCTGAAGGTGCAAGCTCGCG
CATTTCCATCTTCCGCATCGCTGGCCGAGGTACCCCTGTTGCCAACCCAGCATCAACAGTCTCGGAACGTCCCGCAGC
AACCTGCTCAGCAAGATGCGCAATCGCAGCAGGACTACGACAGCGCCTTCGACGACGACATCCCCATGTA

Protein sequence: (SEQ ID NO: 196)

MATPVFWEANIGSAPEHRSPFNNGNPPRQLRLNVMFDNSIPDGQGGYKDRGGFWCSVEWWHQDAQRFALFTKGMVRKV
EGRAIMDRWPKDESSEEVQALKVEASRISILPHRLAEVLLPTQHQQSRNVPQPPAQQDAQSQQDYDSAFDDDDIPM

RL096

DNA sequence (SEQ ID NO: 75)

ATGCGGCAGCTCGATAAGGACCAGCAAGGCGCTCTGGAACAAAGTGCCTTCGCCCACTGCAACAAATGCCTTCCAGGC
GCTGCAACACAGTGCCTCACTAAAAGGCCTTTTAAAGCCTTTTAAAGGTAATAGGGAGCTGGCCAGTTGGCGGAACAGT
CGCAAGCATGGAGCAGGGATGCTTGAACCTTGCCAGGAGCTGCTGGCCAGGTTTCGTCGCCCACTTCACTCTACTG
CCCACCCGACTCATCGAGCAGCGCACATCCGCCCGCACAACTTTCTCCGCTGGCAGCACATTGCATCCCGCGGATGGG
CGTCGGGGTGTGGACGGAAATGCTGGCCAGGACAAAGACCCCGAATACCTGTGCAAGACCTCTACGAGATGGAGCTGC
AGCGCATCACCTCAACATGCAGATCAGCTGATCCACTCCATCGGCAAGCAGGCCGCCGAGTGGCGGAAAAGATGGGC
CAGGCCGAGGCCGAGTTCATGGGCCGACTGCAGCAGAGCACCACCACCTGA

Protein sequence: (SEQ ID NO: 197)

MRQLDKDQGALEQSAFRPLQQTAFQALQHSASLKGLLKPFKGNRELAQLAEQCEAMEQGLLELAQGLLAQVRRPPTLL
PTRLIEQRTSARTTFLRWQHIASRRMGVGVWTEMLRQDKTPEYLLQDLIYEMELQRLTLNMQISLIHSIGKQAAECAEKMG
QAEAEFMGRLLQSTNNH

RL097

DNA sequence: (SEQ ID NO: 76)

ATGGCTGAAACCCATCGGCTTCAGATCGGCTCTCTCCGAGCGATGTCGCTCTGACGCTTCACACCTATCACGCCGCCG
CATCTGGACCGGGCGGAGAAAGTCCGATGCCAAGCACAGCATCTCGGCCTCTCCGCTTCTGCGCATACGTGAATCGCA
TGCACCGCGGGGAGCAGCAGGACGATCCGTAATCCGACTGGTGGCTGGTTCAGATCGAAGAGAAAGTCCGAGAGCTGCCAA
GCCGCACTCGAGGCCATCGACCGAGCTGGATGACGTATGGCCAAGCTGCCCGGACCTCGATATCTCCGAGAACCT
GTCCGTTACACCGGTCAAGGTCCCGTTGTTTCATCTCCAACCTCTCGGCTTCAAGGCAGTCTATCTCTGACCAACTACG
ACGAACTCGCCCGTCCGAATCTGCTGGCCAGCAGCTCGGGCTGGTTCGCTCGCCGACATGGAGGTCTGGCTCGACGAA
GGTGGCTCGGTGCTGCGAAGCCTGTTTGGTCTGGCCAGAGCTACCACTTCTCGGGCGCCACTCGCGACGACTTCGCCGC
AAACAATGCTCGCGCCGAAGCCGCGCGGAAGATGTACGAGAAGTTCGGCGAGATCCCGCAGGACATCCTGGAGGCACTC
GACGCTCGAACTTCGCTCCGCGATCACCCGGGCGCTCTGACGGTGATGCCGATGATGACGCTGACCGTGTGAACTC
GAGGACTGA

Protein sequence: (SEQ ID NO: 198)

MAETHRLQIGSLRSDVALTLHTYHAARIWTRQKSDAKHSILGLSGFCAYVNRMHRAAQDDPYSDWWLVQIEEKVESQ
AALEAIDQRLDDVMKLPATLDI SENLSVT PVKVP LFI SNPLGFKAVYLLTNYDELARRI LLAQHVGLVGRDMEVWLDE
GASVLRSLFGLAQSYQFSGATRDDFAANNARAEEARKMYEFGEIPQDILEGTRRSNFAPPITRGRSDGDADDDADRVEL
ED

RL098

DNA sequence: (SEQ ID NO: 77)

ATGTTCTTGAGCATGGCCCTTTCTTTTGGTCTGCTCTCTTCTGCACTTTTACAGATGCGTGGAACGACCGAGA
ACTCAGGCTGTTGTTAATGCTGATCGTGTTCGGGTATTAGTAACCGTGTGACCATACGGTTGAGATGTATCGCTTTG
AAATGGCGGAAAAGCGATGTGGGAGCTTTATGCAACAAAGCCAACATACATGAACCTGCCAACGATTACCAACGGTAG

Protein sequence: (SEQ ID NO: 199)

MFLSMAPFFLVVLVLSALFTDAWNRDLRLMLLIVFGYSVTVLITIVEMYRFEMAEKAMWGALCNKANYMNCQPDYQR.

RL099

DNA sequence: (SEQ ID NO: 78)

ATGAGAAAGTCTCGGTCGGGCGTCGTGTTTTTGGTGATGCGGCCCGCATCACTCTCCAGGTCCTGACCTCCGCGCCGC
CGGCGAGCTGGGTGATTCCACTGGAATCACTCCACAGGAGCCGACCTCCGCGCCGCCGCGAGCTGGGTGATTCCACTG
GAATCACTCTGCCAGGATCCACTTCGGTATCGGCGGCAAGATGGGTGTTTCGGGCCGAAACACTTCGCCAAAGCGAGGC
ATCACCACCTACGAGGAACTCAAACAATGTTCTTGA

FIGURE 30Q

Protein sequence: (SEQ ID NO: 200)

MRKRSRGVVFFGDAAARITLPGPDLRAAGELGDSGITPPGADLRAAGELGDSGITLPGIHFGIGGKMGVSGRNTSPKRG
ITTHEELKQCS.

FIGURE

RL100

DNA sequence (SEQ ID NO: 79)

ATGAGGCTGTGCGCGCTTTCCCATTTTCGACACTGCTGGACTCGGCCTCGGGGCATCTCGAGGCCCATTTGTATAAGAAGCG
GCTTGCTGCGGAAAGCGGCAACCGCTGGCTCAACAATATTCGGGCATCATTTTCAGCGGCAATCCTCATGAAACCGTTC
CACGGCGCCTCCTCTGGATAAGCGTCTTACTCCGCTGGAGCGGAACCTGCTGGCAAGTTTCCGCTTGCTCATCAACGAC
GACGACTGACCGCGTTCCCGACATACGAGCAACTGCGCCCTATCTCGGTATGACGCGGGCAAGATCGCCTCGCGCGA
AACCATCGCCAAGGCACTCACGGTCTTCTGCTGACCCGCTGGCTCAGCCTCGGCGGACGCTGCGCAACGACCTCAACG
GACAGGTCCAGGGCAACGTTTACATCCTTCACGACGAGCCTGTCTCTCCAGCCGAAGCCTTGGAGCTGGACACCGACTAC
ATGCAAGTTGCTGAGCCAATCCACCGTCCAGGCAACCGAGCCATACGCGAAATCGGGCAGATCATCTGGCGGGAGTTTCTAG
GGATGATCCGGACGTGGGTCGCGCCTCCTACCCATCTGGAGAAGCTCGAGGGACGCTTGAACCAACAGCAATGGGCTA
TCGATAGTCAGCTCGAAGCGGATCCAGCGGCAGAGTTCCGGCATCCGAACCTGTCTCGGATTACCTCATTCACCCCGAGT
TCGGATGCCGAACCTCAGTGAATCAGCGGCAAGCAATGCGCTCTACCGCTGAGTTCCGATACCGAACCCCGACAGAATCC
GCCGAGTACGCCCTTGGTTCCGATGCCGAACCTCATATAGTACGTATACATACAAACAAGATTCTGTATGTAAAAAGCCAG
TACAACCGCGAGCACGCGAGGAAGCCCATCCGAACCTGGCAGGATCTCTGACGCACTGGAGGCCGAGCAACGGATCCAG
GCAGTAAGCGCGCTCAGACGGGTGTCCGAGGATCTTCCGCTACCCATCATCGAGCAGTGGCAGCACCGTTGTGCGGCGG
AACAGTCAGCAATCCGTTCCGCTACCTCATGACGCTCATCCAGCGTCAGTCCAGGGCAAGTTCAACGCTTCTTGGGCTC
CGGAAGAACCGCTGAGCGAACCATCCCGCAACGGAACGCCCATTCGTGCTCCGGCACCATCAAGCCCATAGCGCCT
ACACAGCCTCAGGTCCAGCCCGGGGGATACCCGGACAGGGAGCGAGGTCTCAGCCGGCTCAAGGACCTCATTCCGGCC
CAGGCACGGATCGAGCGTGCCATCCGAGCGGGGTGATGATTCATGA

Protein sequence: (SEQ ID NO: 201)

MRLSRFPITLLDSASGHLEAHLYKKRLAAESGEPLAQYSGIIFSGNPHTVPRRLLLDKRLTPLERNCWQVFRLLIND
DGLTAFPTYEQRLPYLGMQPGKIASRETIKALTIVLRRLRWLSLGRRLRNDLNGQVQGNVYILHDEFPVSPAEALELDITY
MQLLSQSTGHGNRAIREIGQIIWREFRDDPDVGRRLPTHLEKLEGRNLHQWAIDSQLEADPAAEFGIRTLSDLPHSTPS
SDAELSEISGKCALPLSSDTEPRQNPSTPLVRMPNSYSTYTKQDSVCKKPVQPRAREEAHPNWQDLHLHALEAEQRIQ
AVSALRRVSEDLRLPIIEQWQHRCAAGTVSNPFGYLMTLIQRAVQGFNASWAPEEPAERTIPATERPIRAPAPSSPIAP
TQPQVQPRGDRTRTGSEVLRLKDLIRPRHGSSVPSEGGDS.

RL101

DNA sequence: (SEQ ID NO: 80)

ATGTCGAAGTCGACGATCAATGAAGCGGTCTTGACGCAGGTGCTCAACCACCTGCGCAACGGCCAGCTCAGGCGTTGTGC
CGAGATGGGGCTGCGGCCGAGATTCTGGCTCAGCTCCAACAGCCTGCCGTCATGAGCATCTGACCAATACCCCGGTTT
CCTGGGTAGATGTGAGTGAACATCGAGCTCATGGAGAAAATCCTGGCCACAGCCGAGCGCAGCGCGAGGAAGACCTG
CAGATCGAACCGCACTGAGCTGGGAGCCACCACAACGATGATCCAGAGCTTTTTCGGTCTGTGCGCGGAGGACACCGC
CACCAAGCGCTTGATGCTGGAGATCCACCGCGCCGCGGTGCTGGCGGCAGCTCGATGAACAGATCGAGCGCCAGATAT
GGTTCCGCTGGGAGCACCTGATGAGGAAAATCAGGTCGCGCTTGAAGACAGCATGGAGTTGCTGGACATCGCGATGATC
CTCACAGAGGAAATCAACGCGGAATCGAACAAGACAGTCCAGAATTCATCAGCCTCGCCATTGTTTGGTCTCTCATCCA
GAGCTGGTTGAAAGACGGGCTCTATCCGTCTGGCAATCGAGCCAGAGCCAGGCGGGCTGCAAAAGTCCCAATCCACTC
TTTACCTCGCTAGCGTCAGCTCACACCTGCCCACTCTGCCCATCCGCAACAACGAGGTGAACGCTGAGACAGAACGT
CAACAACCTACTGAACCTGGTTAGTCGGAAGCGACACAGCACCATGA

Protein sequence: (SEQ ID NO: 202)

MSKSTINEAVLTQVLNHLRNGQLRRCAEMGLRPEILAQLQPPAVMSILTNTPVSWVDVRVNIDVMEKILATAERSAQEDL
QIERALKLGATTTMIQSFFGLSPEDTATKRLMLEIHPRRGRWRQLDEQIERQIWFRWEHLMQENQVRLEDSMELLDIAMI
LTEEINAGIEQDSPEFISLAIVWSLIQSWLKDGLYPSGSSQAGLQKSQSTLYLASVSSHLPSPSATTQVNAETER
QQLLNLVQSEGDTAP.

FIGURE 30R

RL102

DNA sequence: (SEQ ID NO: 81)

ATGAGTATGGCCAAGATCAACCCGCAAGATCTGAAGACCGGCTACTTGCCCTGGTTTTACCGCACCGCCCAAGGTTCT
GGAGCAGTTGTCGGACCCGATCAGCGACACGCCCATGAGGCTGACACTCCACGACGTCCTCCCTGGCAGGATAACCCCC
GGACCACTCGAAACCCGAAATACGATGAGCTGAAAGAATCGATCCGACATCGAGGCTCGATACGCGCCGACCACTGACT
CGACGCCCTGGAGAGGACAAATACCGCATTCGCAACGGCGGAAACACTCGCCTGGAAATTCTCAACGACCTCTACAAAGA
GACCGGAGACGAGCGCTATTTAGCTTTCAGCTTTCAGCTGCGCTGTTCAAGCCCTGGGACAAGCAGCGCGCGGAAATCATCGCGCTGA
CCGGTCATTTGGCCGAGAAGCATCTGAAGGGCGACCTCAAGTTCATCGAGCGCGCGGTTGGGGTGCGAGAAGGCGAAATTT
CTTTACGAACAAGAGAACGGCGGTGAAAGCATTTCCAGCGCGAGTTGGCACGTCGGCTAAAAGCGGACCGCTACCTCTGT
ATCTCAATCCCATATCAGTAAGATGTTAGACACTATTGAGGTATTGGCGCCGCGGATTCTGTGATGCTGTATTAGGGC
TCGGTAAACCGCAAACTCGAGAACTCCTGTCACTCAGAAAGTCGGCATCCTCCTGCTGGGCGACGTCATACGCTGGTGAA
GGGGTTGACTTCGAAATGCTGTTCCAGGACACCTTGGCAATCTTCGACAGTAGCCCTGACGAATTCATTTTCGAGCGTTT
CCAGGACGAACTCATCGACCAATGAAGCGCCCTGGGCTGCGTTATGACCAAATCCTGCTCGAGATTACCAACGGCG
AGCAGGAGCAACGCCGCGCACTCTGGTCGACCTGCCACACCTGCCGACCACTCAACTCCCAATTTGGGCGAGGAA
AACCCTGCTGCGTCGTCCTACTGGACAAGCACAACACAGAGCCCGCCGAGATCCCAAACGTCAGGACAAGGAGCAA
CCCGGGAATCCCTCCCGCCGCGCTCCGCCACCACTGTCCAACAAAAGCAATTGCCGATGAGGAGCGTGCGGCGG
TCTTGGCAGGCCATATCGTGAGCCCGGTATCGACTAAGATCCAGCAGACTCGCAACGGCTGGCCGCGCTCGAGGGGAA
CATCTACCTGTCTTCGATGAACAGCTCTGCAGGCAATCCAGTGCAAGTCGGTGGCTGCACCCGATCACCATCTCTG
GTACATCGAGGGTTCGATCGATACCCCGAGATCCTGCGACAGACATCGCTGATCTGGGTGAAGAGATCGCTCTGCATG
TCGGCGCCCGCAGGCGAGATCGTCAGGATTAGGGCGGTGTGGGTACACGTATCGCGAGCCCAATGAAGACCATGAGATT
ACTGATTCAGCGCTCGCACTCATGACGCTTCAAGCGCTCAGCGGCGAGGTCCAAGTCGTTCTGAACACTCAACGATCA
ACAGACCTGCCGCGATGCACTGGGTGAATTCAGTTCTCAGCTGGCCTCGCTCAGTTGCTGCTGGGCAACCCACCACAA
GTGACAAGCCATCCTGCCAGGCGGCGTCTCAATGACGAAGCCCTGGTGAAACTGTTCCGGATCATTCGCTTGGCCGA
CGCTGGTTGACCTTGAGCTGCCCGCGCGCTCCGAGCAAGCAGCTACTGACCAGTGA

Protein sequence: (SEQ ID NO: 203)

MSMAKINPQDLKDRLLAPGFTAPPKVLEQLSDPISDTPMRLTLHDVLPWHDNPRTRNPKYDELKESIRHRLDTPPPVT
RRPGEDKYRIRNGGNTRLEILNDLYKETGDERYFSDFCLFPWDKQGEIIALTGHLENDLKGDLPFIERAVGVQKAKF
LYEQENGGESISQRELARRLKADGYPVSQSHISKMLDTIEVLAPAI PVMLYSLGKPKQIEKLLSLRKSASSCWARYAGE
GVDFEMFLQDTLAIFDSSPDEFIFERFQDELIDQMKRPLGLRYDQILLEITNGQQEQRRGTLVDLPTFAAPPQLPPIGQE
NPAASSTGQAQTQSPAADPQTSRTRSNPGNPLPPPPPPVQKQLPDEERAALVAGHIVSPVSTKIQQTRQRLAGLEGE
HLPVFDETALQAI PVQVGLHPITDLWYIERSIDTPEILRQHIADLAEELALHVGAPGEIVRIQGGVGYTYREPNEDEHI
TDSALHMTLLQAVSGQVQVVLNTHDQTCRDALGEFQFSAGLAQLLLGQPTTSDKPSQAGRLNDEALVKLFRIIRLAR
RLVDLELPPAASEQAATDQ.

RL103

DNA sequence: (SEQ ID NO: 82)

ATGACCATGGCCCGAGAAACCGAAGATAAGTTCGTTGTCCGTATGCCCTTGGGCTTGCGCGATCAGCTAAAGCAAAAGC
CGCGGATAACACCGTTTCGGCCAACAGCGAGATCGTCTACCGACTGGAGCGCAGCAACGCGCTCGAAGAAGAACTCGCGC
GAGCAACCGAATGGTCGACGAATCTTCGCCAAGAACCAGCGCTGCAGGCTGAGCTGGCGGCGGCAACACGCTCAG
GTGGCGGAGGCATGA

Protein sequence: (SEQ ID NO: 204)

MTMARETEKDFVVRMPLGLRDQLKQKAADNHRSEIVYRLERSNALEELARANRMVDELFAKNQRLQAEALAAANTPO
VAEA

RL104

DNA sequence: (SEQ ID NO: 83)

ATGCCTATCAAACACGCCATCGTCCACCTGATCGAGAAGAAGCCCGACGGCACCCCGGCGTGCTCCACGCGCGGACGCG
CGAGCTGGGCGACTCCCAGGCCATCGAGAACCTGCTGGCCGACCTCAACGAAAGCTACAACGCCAAGAAAGGCGCTGGG
GCTTCTTCAGGGCGAGTCCGGGGCTACCCGTTACAGCGCTGGCTCGGCGAGTACCTGGAGGGCGACCGGACTTCGTC
GGCTTCAGCGCGGAAGCGGTTCGAGCACTGCAAAAGCTGATGGAGGAGTCCAATCTCTTACCGCGCGCCACGTCCTGTT
CGCCCACTACCAGCAAGGCATGACCGACTACCTGGCGATCGCCCTGCTGCACCACAGCGAAGGCGTGGCGGTGAACGAGT
CGCTGGAGGTCAACCCGTCGCGCCACCTGGACCTCGGCCAGTTGCACCTGGCCGCGCGGATCAACATTTCCGAATGGCGC
AACAACAAGCAGTCGAAGCAGTACATCTCGTTTCATCAAGGGCAAGGGCGGGAGGAAGGTCTCCGACTATTTCCGCGACTT
CATCGGCTGCCAGGAAGGGTGGATTTCGCCGAGCGAGACGCGCACCTGCTGAAAGCCTTCAGCGATTTCGTGGAAGCG
AGGACATGGCCGAGGAACAGGCCCGGAGAAGACCGAGACGCTGGTCGACTACGCCACCTCGCAGGCGCGCATCGGCGAG
CCGATGACCTCGACGCGCTTTTCGAACTGATGGACGACGCAACCGCGGGCGTTCTACGACTACATCCGTAAACAAGGA
CTACGGCTGTTCGCGGAAATCCCGCGGACAAGCGCACCTCAACCAAGTTCCGCGCTTACCGGCGCGCGCAAGGCC
TGTGATCAGCTTCAGGCGCACTGCTGGGCTCCAGGATCGAGTACGACGAGGAGCGGACACGCTGCAGATCAGCAGC
CTCCCACTCAACTCCGCGACCGCTCAAGCGCGCAAGGCCCAAATTGGAGAATGA

FIGURE 30S

Protein sequence: (SEQ ID NO: 205)

MPIKHAIVHLIEKKPDGTPAVLHARDAELGDSQAIENLLADLNESYNKKNKAWGFFQGESGAYPFSWLGLEYLEGDRDFV
GFSREAVEHLQKLMEESENLF TGGHVLFAHYQQGMTDYLAIALHHSSEGVAVNESLEVTPSRHLDLGQLHLAARINISEWR
NNKQSKQYISFIKGKGGRKVSDFRDFIGCQEGVDSPSETRTLKAFSDFVESEDMAEEQAREKTETLVYATSQARIGE
PMTLDALSELMDQPPAFYDYIRNKDYGLSPEIPADKRTLNQFRFTGRAEGLSISFEAHLGSRIEYDEERDTLQISS
LPTQLRDQLKRRKAQIGE.

RL105

DNA sequence: (SEQ ID NO: 84)

ATGCGTAGTTTCCTTCGCGGCGCCCGGAAAGCGTTCGCGCGCTGGTGGCCTTCGCTCAAGCAGAAGGCTGGAGCGTCGA
CCGCTCCGAGGCGGCCACTTGAAGCTCAGCAAGATCGGCTGCGCCTCGATCTTCATTCTTCCACGCCAAGCGACGCAC
CGGCGAGCTCAATGCCGCGCCCTGCTCCGTCGAGCCGACAGGCAGCGTTCCTGAACCAGGAGTCTTCTGA

Protein sequence: (SEQ ID NO: 206)

MRSFLRGARESVRRLVAFQAEGWSVDRSAGGHLKLSKIGCASIFISSTPSDARGELNARALLRRADRQRSLNQESF

RL106

DNA Sequence: (SEQ ID NO: 85)

ATGCGTGACGTCACAGCCTACCGGCGCTCGAGCACTTCAGAAAGTCGAGCTGATGCTTGAGCTCAAGTTGCGTGAAGG
TCCTTCGTGGATCTGTCTGAAGTCCGGCTATCACTGGATGGCAGCGGCGCACAGCCCTGCCCTGACTGCGGAAAGTCGC
GCTACTGGACACAGCGGTTGGAGTGTAGGTCGTGGCCATCGCTTCTCGGCAGCAAGGGAAGAGTGGGAAAACCGCCTCAGG
ACACGGTCGCGGTACCTGTGCGCTCAACGGCACCAGTAGCAACTGACGACGTATGCACTCAACTGCGCACAGAGGTCCG
CATGCTGCGTTCCGCGCATGACGACCTGGCCTGCAGCCGCGAGCGATCGTCGAGCCTTCAGGCGCTGGTGAAACGTC
TCCTGGATGCCGCCGCCACCGATAGCCTTCCCCGCTCCCTTGAGAGATGGAGACCTGGCTGCAGCTCAACAGCGAGGAG
ACCACGAATCGTAG

Protein sequence: (SEQ ID NO: 207)

MPDVTAYRPLEHFQKVELMLELKLREGPSWICLNCGYHLDGSGAQPCPCGKSRYWTSGWSVGRGHRFSAAREEWENRLR
TRSRSPVASTAPVATDDVCTQLRTEVRMLRSAHDDLACSRQSDRRSLQALVKRLLDAAATDSLPRSLAEMETWLQLNSEE
TTNA.

RL107

DNA sequence: (SEQ ID NO: 86)

ATGAAGGCGTCCCAGACCTATCAGTGCATCGTCAAGTTGATGGCGCCGTTTCTGGACCAATACCATTAGAAGCAGCG
TGCGACCTGCACCTGGAGCGACAAGGTGGCAGCCTCCCGCCTTGCCGAACGACTGTTTGGCGAGGACAACGCATACATCA
CCCGTATGCCGGTACAGGCAGGCGACCACGAAAAGCGCATCGAGAGCCGCTGGGCGCTGTCTGTAGAAATCCCAAGGAG
GTAGCGCGCGATGCCTGA

Protein sequence (SEQ ID NO: 208)

MKASQTYQCIVKFDGAGFWNTIQKQRATCTWSDKVAASRLAERLFGEDNAYITRMPVQAGDHEKRIESRWALSCRNPKE
VARDA

RL108

DNA sequence: (SEQ ID NO: 87)

ATGAACACTGAAGCCCGCTTCCGAGTATCCACGCCTCGGCCGCTTACCGACTCGGCAGTGGTTCATGCCAATCACGT
TGGGGTCAACCCCATCGAGCTGGACGCCCTCAGCCAAGTGATCTCGCGCCTTTCGCGGGACGAGAGCACGGTCGCACCCA
GTTTCGATGGAGCGAGAGCTTCGTGAGCTGGAGGAAGTGGGTACATCGAAATCTCGACCACCCAGGCCGGGACTCTGGTG
GTCACACGCGCGCTCCGGGGCAATTGCTTTCCGGCTTACTTCTGGTCCGTATGGATCCCGCGACACCTGTTAGCTGCTC
GCTGAAAGTGAGCCTGGTGCCGACCTCTGCTGCGGCACTCAGGACTCCAGCACCTCACCGCCGTGTTCCGCATTGCAG
GCAGCAAGGACGCCGCGCGAGTTCCTGCATCAGTTGGCCAACTATCCCGGGCATGAGCCGGAGTTGCCCGAACTG
GTGCGCGTTCAGGTCGGTGATGCACTCAGCAAGGAGGCCGAGTCATGA

Protein sequence: (SEQ ID NO: 209)

MNTEARFPSIHASAAFTDSAVVHANHVGNPIELDALSQVISRSLRDESTVAPSSMERELRELEELGYIEISTTQAGTLV
VTTRAPGQLLSAYFWSVWIPRHLFSCSLKVSIVPHLCCTQDSQHLTAVFRIAGSKDAAREFLHQLANNYPGHEPELPEL
VAVQVGDALSKAEAS.

FIGURE 30T

RL109

DNA sequence: (SEQ ID NO: 88)

GTTTCTGCAGATCAGCCGGGAGGACTCGAAGTGAGTGTCTGGAAGTACGCCGCCGCACTCCGTCGAGGCGGAGCAAGG
GGTACTCGCGCGCCTGATGCTGGACAACGCCGCATGGGACATTGTCGGCGATCAGTTGCAGAGGAGGACTTCTTCGGC
ATGAGCATCGGCTGATCTTACCGCCATCAGCGAGTTGGCGCGAAGGATGCTCCGTTGATGTCTGTAAGTGTGTCGGAA
GCGATCGAAGACCTTCCAGAAGCTGGCGGGCTGGCTACCTCGGCCAGCTCGCCGACAACACGCCCTCCGTGGCCAAATAT
CGAGGCTTACGCGCAGATCGTTCGCGATCGGGCACCTCGCGCAGCTGATGTCTCTCGGGCACCCTGCACCAGGACCG
CCTCGAACCACAGGCAATCCCTCTGAGGTTTCAAGAGGAGATTGAGCAGAAGCTGTTGGGCCCTTGGCCAGGACCACCA
CAACGCCGATTTCGTGATATCAACAAGAGTCTCACGAAGATCGTTCGACACCATCGATTACCGCTTCAACAACAACGTGA
CGGTAACGGGGGTCGCCAGTGGCCTGAAGGATCTCGACGCACTCACCGCGGACTACAGAAGTCCGATCTCATCATCGTC
GGTGCCCGCCCGCGATGGGCAAAAGCTCGTTTGCCTCAACCTGGTCGACACCGCGCTCCAGAGCGACCAACAGAAGTC
TGTTCAAGGTGTACAGCATGGAGATGCCCGCAGAGCAGTTGCTGTTCAGGCTTGGCGCCCTGTTGGGCCACCTGGACCTGG
GCAAGCTGATGAAGGGCCAACTGCAAGAAGAGGATTTGGCCAGACTGTCTGGCGCGATCCAGCGCATAAACGACTATGGC
AGCCGGCTGGTCATCAACGATCAGGGCAACCTCACGCCGACAGCTGCGCGCCAAGGTTCCGCGGGCGGCGGAGGTA
CGGACACCCCGCGCTGATATTGGTCGACTACCTGCAACTGATGAGTTGCCAGGCTGGAGAATCGAGCCACCGAGATCT
CGGAAATCTCCGCTCGCTGAAAGCGCTGGCCAAGGAGATGGACTGTCCCGTCGTAGCTCTATCCAGCTAAATCGCGGC
CTAGAGAACCGGACGAACAAGCGACCGAACTGCGCGGACCTACGAGAGAGCGGCGCAATCGAGCAGGACGCGGACGTGAT
CATGTTCTGTGTACCGCGACGAGGTCTACCAACCCCAACACCGAGGCCAAGGGCATCGCCGAAATCATCATCGGCAAGTATC
GCAACGGTCCGATCGGCACGCTCCACACCGCCTTCATCGCCAACAGACCCGCTTTGCCGACCTGGCGCGGGGACCTGG
CAA

RL110

DNA sequence: (SEQ ID NO: 89)

ATGACTCGCTCTGCTCTCTCGACCATCGCCTACGAGGCCCTGGTGGTGGCCCGCAAATTCAGCAACCGAGAGGAGCG
CTGCATCCGCGAAACCTGGACCGCCGAACAGGAACTGGTGCTGCTGCGCCTGTATCCGGATATGCCGAACGAGGTCTTGG
CAGCCAGGTTGAACAAAACGCTCCAGCAGATCTGCTCCAGAGCGTATCGGCTCGGGCTGAAAAAAGCCCTGAGTTCTCC
AAGAAGATCCGCGAGGACTGGGGCAGCGCAACTCGGTTCAAGAAGGGAAACACCCCATGGAACTGCGGCATGAAGGGGCT
GCCCGCGCAGGACCGCACCAAGAAACGAGTTCAAGAAGGGGCAAAAGCCCCACACATGGCTCCAGTCCGCGAGCAGCG
GGGTGAGCGCTGATGGCTACCTGCAACGAAAGATCTCGGATACCGGCTATCCCCCGGGGACTGGAAGGGCATCCACATC
CTGCTCTGGGAAGAACACTTCGGCCCCATCCCAACCGGCCATTGCGTCTGCTTCAAGGACAACAACAGCAGAACGTCGT
CATCGACAACCTGGAGTCTATACCCGGGCCGAACGCATGCGCCGCAACTCCATCCATCGCTATCCACCTGAGCTGAAGA
GCGCAATCCGCGTATCAGCAAGCTCAAACGCACCATTCAGGAGGTCGAGCATGAAGAACAAGATTGA

Protein sequence: (SEQ ID NO: 210)

MTRSLSTIAYEALVRARKFSNREERICIRETWAEQELVLLRLYPDMPNEVLAARLNKTLQOICSRAYRLGLKKSPEFS
KKIRQDWGSATRFKKGNTPWNCGMKGLPARGRAPETQPKKGQKPHWTLPVGSSTRVSADGYLQRKISDTGYPPRDWKGIHI
LLWEEHFGPIPTGHCVCFKDNKQNVVIDNLELITRAERMRRNSIHRYPPELKSARIRVSKLKRTIQEVEHEEQD

RL111

DNA sequence: (SEQ ID NO: 90)

ATGGACAAGCAAAAAGTCTCTCGCAAGGTCGAGAAGCTGATGGCCCTGGCGAATGCCAAGGGGGCCACGCCGAACGAGGC
GGAACCGCATTGCGCCAGGCCGCGATCCTGAAGCGGCAGTTGACCTCAGCGATGCGGAGATCTCGGCCCAACAGGTGG
AAACCGCGTGGTTCCTCACTCGAACCAGGCGCTCTCTGCCCCATGGCTGCATGAAGTGGCCGGGATCTGCGCCAGTTCC
TTGGCTGCGACTACCTGGCGGCATACGCGATGCCAGCGGGCTGGACGTTCAAGTTTCATGGGCCGAGGGATCGGCCCTGA
GCTGGCCGCTCAGCCTACTCTACGCTCCACCACCAACTGGTGGCAGCGCGCTCGGCTCATGTCGCCAACAGAAGCGCT
GCAAGCTGTGACCAAGCGTCGTGCGAGCAAGCTCTTCGTGCAAGGCTGGCTTCTCGCAGTGCCTTCGTTGACGTGAA
TTTGCTGGCAGGCGGACGAGTCAAGCAGCCATCAAGGCTACCTCGAAGTACACCATCCGGCGTTGAAGTACCT
GGAGCGGCGCGCTTACGAAGGCCCTTGCTATGACAGGCGCTCGCTGCAAGCAGGCTGGGAGCACGGCAAAAACACTC
GCCTGCACCGCGGTGTGACCGCGGAGTTCAGGCGCGCTCGAGCAGGAGGTTCCCAATGA

Protein sequence: (SEQ ID NO: 211)

MDKQKVLAKVEKLMALANAKGATPNEAETALRQAAILKRQFDLSDAEISAHTVETACVPTRTRRSPAPWLHELAGICASS
FGCDYLAAYAMPAGWTFKFMGRGIGPELAHAAYSTLHHQLVAARSAHVAQQKRCKLSTKRRRSKLFVEGWLLAVRSLVRE
FAGRPDESTQAIAKAYLELHHLPALKYLEPAALTALAYDQASLQAGWEHGKNTLRHGRVSRVQGALEQGGSGQ

FIGURE 30U

RL112

DNA sequence: (SEQ ID NO: 91)

ATGAGTGACCCCAAGCTCAAGCCCTGCCCGCTCTGCGGCAGCACGAACATTGGAATGCTGGAACCCGAGCTGCTCGACAC
 CGATGCCTGGAACTGTGCCATTGAATGCCTGGACTGCCAGGTTACATCGGGCCGCTCTACTGCGAGCCAGACCCGGTAA
 CAGCGAGGTATTACGACAGATCGACTGGAATAGACGCCCAAGCGCAAAAAACACGCGGACGAGCGTGAGCAGTTCTTG
 ATGGCCAACTGCTCGCCGCTGGAGGTCGCACTGGGCGACGTAGCAGCCCTGGCTATTGTCGATCGGGTAAGACAGGC
 CACAGACCGAATTTACCCAATCTCGAACCTCTCCCTGTTCGCGAGGCTGGCTCGATGTACAGGCCGAGCGCCGCGCC
 AGATCACCGTCGAAGGTTTCGATACAGCAACGACGCTAGCGCTGGCCTGATCGCCCTGGCGGCCGCTGCTACGCG
 CTCCATGCGCGCGCATCGGCACCGACTGGCCGGCGCGCATTCGGAATGGCTCTGCACTGTTCTGGCCCTGGGACGAAGA
 GTGGTGAAGCCTAAGTCGGCGCGCGAGAACCTGCTACGCGCGCGCCCTAGTGTGGCCGAGATCGAGCGCTGGACC
 GCTCCGCCACCGAGCAGGGCTCAACCATCTGCAAGGGGGCGCGTAA

Protein sequence: (SEQ ID NO: 212)

MSDPKLKPCPLCGSTNIRMLEPELDTDAWNCAIECLDCQVHIGPSYCEPDPVTARYSAQIDWNRPSAKNHADEREQFL
 MANLLAALEVALGDVAALAIIVDRVRQATDRIYPTSNLSPVPQAWLDVQAERRRQITVEGFDTSNDDASAGLIAALAGCYA
 LHAGGIGTDWPGGIRNGSALFWPWDEEWKPKSARENLVRAALVLAIEIRLDRSATEQGSTICKGGA

RL113

DNA sequence: (SEQ ID NO: 92)

ATGAACCTCCAGAACCGAACAACCTCCTACTGAGCTTGATCGCCGAGACCCAGTTGACGCGCTACGTGCAAGGCTACAT
 GGCCAAAGCAGGCGCTGCCCGGCTGCTTCCGAGAATCTGCAATCGAGGCTGAAGGTGCTGCGATGTTGCAGGGCCTGG
 TCGCTCCGGTTTCGCGCTCAGCAGCGTGCTGTGGACAGTCCCTGCGAAGCGCACTGCTCCAAATCGCCACGACCTACTG
 TTGCAGACGAAATCGCAACTGGCCATCGCGGCCAATGCCAGTTTCGATCCAAAGTATCCAGCGGGACATGAACAGGGCGAT
 CTGGAACATAGCTACTGCCATCGATCACCTGGCCGAGTTTCGCCCAACCTCGCAGGACACTGTGAGGGTCATCGAACGGC
 TGATGCTCTTCGTCGCGCAGCTCATCAAGCACTGAAGGCCAGCACTGGCCGCCGAGGCAATGCGGTGCTCGGCATGAGC
 GTGGGAGGCGCTGGCATGA

Protein sequence: (SEQ ID NO: 213)

MNLQNRNLLLSLIAETQFDAYVQGYMAKAGAAAGASENLQIEAEGAAMLQGLVAPVRAQQRACGQSLQNALQLIAHDDL
 LQTKSQLAIANASSIQVIQRDMNRAIWNIATAIDHLAEFAQPSQDTRVRIERLMLFVGSSTEGQQLAAEANAVLGMS
 VGGLA

RL114

DNA sequence: (SEQ ID NO: 93)

CTGAACAAGTTTCGCGAGCGCCGCCACCTTCGGAGCCAGCAGGCCAAATTGACCGGCGCTACGCGAGAAATACGCAAGCT
 GACTGGTGGCGGTATCGACCTGTTTCGGGAAGCTGGGTTGCTACTTGAGCTTCGAACAAAAGCAGCTCCTACAAGACGCAG
 CGCGCTTGCTCGACTCGGTGAACAAGCAGATCGAGCATGCGAAGGAAAAGCGTGATCGCTACGAGAAAAAGCCAGAAG
 CGGCGCAGCTACGTGAGCGCCTGGCCAAAGCAACTGGTTCGCTCGCACTACCCGCTTCGGGAAATACGCTCGAAGATCG
 GCTGGAATCCTGCAGATCGCGTTGATCTACAACCGGGCCAGGGTGTTCGATCACCTGTACTCCACGCACCAGCTCCACT
 CAAACTCAAACGCTGGCTGGAGCGTCCAAGCAGCTCATCGGATGGCGCAGTGAAGCCGAGTATTTCGCTAGTCAGGTG
 GGGAGCCTGCGATGTGACTTATTAGCCATCTGACTAACGAAATCGCGTACGACGATGGCAGTGAAGTCGAGGAGCGCCT
 GCGCGTCATCAAGCAGAAGGTCGCTGACTGCACCGCACAGATCGCTTGACCAGCGAGGAGCAGGAAACCTTCGGCTCT
 GGACAGACGCTCTGCAATCGGCTCCGGAGGGCCTCATATGA

Protein sequence: (SEQ ID NO: 214)

LNKFGSAADLRSSQAKLTGATREIRKLTGGGIDLFGLKGCYLSFEQKQLLDQAARLLDSVKNQIEHAKEKRDREYKAKK
 RRELRLERLAKQLVASNYPLPGNTLEDRLLEILQIALIYNRVRFDHLYSTHQLHSLKRWLERPKQLIGWRSEAEYFASQV
 GSLRCDFISHLTNEIAYDDGSEVEERLRVIKQKVADCTAQIALTSEEQETLRLWTDALQSAPEGLI

RL115

DNA sequence: (SEQ ID NO: 94)

ATGAATGCGAAAGCGACTTCGGTTGTATCCACCAAGGTTGGTGTAGGAAAATCCACCACCGCCGCCAACCTCGGTGCATT
 TTGCGCCGATGCAGGCATACGAACCTCCTCATCGATCTGGACCCCGTCCAGCCCTCCCTATCCTCGTACTACGAGCTGC
 CGGAAGTTGCCCAGGGCGGCATTACGACCTGCTCGCGCCCAACATAACGGAACCGGCGAGGATCATCTCCAGGACGATT
 ATCCCCAATCTGGACGTCGTGATTTCCAACGACCAAGCAATCAGCTCAACAACCTACTGCTCCAGGCGCCGATGGCCG
 GCTACGCTGGCGAAGCTGATGCCCCGCTCTGAAAGAGGCTACGACCTGGTGTGATCGACACCCAGGGTGGCGGCTCAG
 CTTTGCTCGAAATGGTTGTGCTTGATCGGACCTGGTTGTTTCCCCCTCCAACCAACATGCTTACCGCCCGTGAGTTC
 AACCGCGGACCATGCAATGTGACGGCCTACGCCCTATGAGCGTCTCGGCATGCGGATCCCAATGTTTCAGATCGT
 CATCAATGCTGACCGAGCAATGACTCCCGGGCAATTACGAGAATGTGCGTGCCATCTTCGATGAGCATCAGGACA
 TTTCTGTGCTCGAAACGATGTCCCGGATGCCGTCGTGTTTCGCAACGACGATCGCGCGGGCTACCAGCGCACCGCCTC
 GAAACGCGGCAACCTCCAATCGCACATCAGCGCCGCGCTGGAAATCATTGAAACCTGGCCATCGAGGTCTTTCCGA
 GTGGACTGACCGCTTCTGGCGCTGACGCGGGAGGCGGTTCAGCACTGGTCAAGGGAGGGCGCTGACATGGCGAAGAC
 TCCTATACCCAAGCCCGCAGCTCGACGCGGAACCTGTGCTGGAACCTGA

FIGURE 30V

Protein sequence: (SEQ ID NO: 215)

MNAKATSVVSTKGGVGKSTTAANLGAFCADAGIRTLIDLDPVQPSLSSYYELPEVAQGGIYDLLAANITDPARIISRTI
IPNLDVVISNDQNNQLNNLLQAPDGRRLRLANLMPALKEGYDLVLIDTQGARSALLEMVVLASDLVVSPLQPNMLTAREF
NRGTMQMLDGLRPYERLGMRIPNVQIVINCLDQTNDSTRAIHENVRAIFDEHQDISVLETTVPDAVVFNRNAASRGLPAHRL
ETRQPSNRTSAPALEIIRNLAIEVFPEWTDRLFALTPGGGCSTGQGRALTWRRLSPKPATSTRNLCWN

FIGURE 31A

RS01

DNA sequence: (SEQ ID NO: 95)

```
ATGGGGATCTACCGCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAGGGGCGGGCTACCGGACGAAAGGTAGCTGCG  
CCTCCACGAGTTCGCTAGGCCTGTAGGAAAAATCTGGAATTACCGAGAGCGCTGGATTCCAGCGCCGGCATGCTGGCA  
GAGCCCCGCAATTTCAAGGCCGAAACCGCAGTACCCCTCTGTAATCGCTGATTACGTCGAGGGCACATTGCTACGCCCTGCA  
GAATGGTTTCAGGGCCTGAAAAACAGAAAGCCACCTAAATAGGCGGGCTATTCCATATTGACATCACGTCAATGCGGG  
CC
```

RS02

DNA sequence: (SEQ ID NO: 96)

```
ATGACGCGGCAGCAGCTACCGAGGAGTACATCTTCGCGCACGATCTCCGAGAAGCCAGCGCGAAGATCTACCGCGCCGC  
GACCAAGGCGCTGCTCAAGCACTTCGGCCCTACGGCAACCGTACAGGAGGTGGACCAAGGTCTGTGCTGGGATGGCGGC  
GCAAGGTCCTTGAACAAGGCTGTGCAAGCGGAGCTGGAACACGTATTCGAATCATCTGCGAACGATCTGGGGCTATGCC  
ATCGAGCAGCAGCTGGTGACACACTCCCAAGTCAACCGTTTCAAGAACACCGTTCATCCCCCAGGCGAGCAAGCAA  
AACCGTCGCAGCCGAAGCCATCCTGCGCGCCCGCAATTGGCTCAACATGCAGGTCGGCGCCGAGCGCTGCACTGGCGATC  
GCGCACGCATCACTCCCGCCTGGTTCTGGCTTTGACAGTTTGGAGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTG  
TTGTGCATCCGCAAGCGGCACATCGACTGGGAAAATCAACTGATCCTCATCCGCGGCGAGACAGAGAAAATCATAAAGA  
GTTCTAGTGCCAATAACGGAGGGGCTTGTGCCTCACCTATCGCGGCTCCTGCAGGAGGCCGATAGAGCCGGATTTCGCCG  
ATGACGACCAAGTTGTTCAACGTCAACCGGTTCTCACCGCACTACAAGAGCAAGGTGATGAACTCCGACCAGGTGGAAGCC  
ATGTACCGGAAGTTGACCGAGAAGGTTGGGGTGGCGATGACTCCGCACCGTTTCCGGCACACCTGGCCACCGACTTGAT  
GAAGGCACCCGAGCGGAACATCCACCTCAGGAAGTGCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCG  
AGGCCGACTACGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAAGCTCAGGAAGGTG  
GATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCAACCTCTCGCTCGAATGGGTGAAGCGCCGCC  
ACAGGAGGCTAGGACAGAACCTGCAGAACCAAGGAGCACACACAGGGACAGGCATTACGGGAGATGCAACCGCGTGGG  
AAGAAGCGCTACACAGCCCTGACACCTTCGAGCAAGCGTGCTGTTCACTCTGATGGCTCAACACCTATCGAACCGT  
GCCGCCACGGCCTCCGCGGCTTCACCGCAACAAGCGGATCTGGAGGATGGGGATCTACCGCCCGAAGCAGTCTCGCCTA  
G
```

Protein sequence: (SEQ ID NO: 216)

```
MTPQQLTEEYIFAHDLREASAKIYRAATKALLKHFGPTATVQEVDRHSVLGWRKRVLEQGLSKRSWNTYSNHLRTIWGYA  
IEHELVTSHSQVNFPRKTTVIIPRRASKTVAAEAILRARNWLNMQVGAERCTGDRARITPAWFWLCTFEVYFTGIRLNLAL  
LCIRKRDIDWENQLILIRGETEKTHKEFVVPITTEGLVPHLSRLQADRAFGADDDQLFNVNRFSPHYKSKVMNSDQVEA  
MYRKLTEKVGVRMTPHRFRHTLATDLMKAPERNIHLTKCLLNHSNIQTMSYIEADYDHMRVLAHARSLAQGALENVRKV  
DYGSPQASAKPKPCGQPLARMGEAPPQEARTPEAPREHTPGTGIQGDATAWEEALPQPPDTFEQSVLFTLMAQHLSNR  
AATASAASTATSGSGGWGSTARSSLA.
```

RS03

DNA sequence: (SEQ ID NO: 97)

```
ATGAAATCTGGTATCGCGACCGTCGCTGTTTCATCAACGACACCAAGGCTTTGGTGCATACCGTCGACGGGACCGCCAT  
GCTGGTCACGCCAGGAATCTTCAAGCGTTATGTCCAGGAGCATCCGGAGGTTGAAAAGCTGGCCCAGGCCAAGGAGACCG  
CCGGCTGGAAGCTGGTGCAGCGCGGTTTCGAGAAACAGGCTTTCACCGAAAGACCAGTAAGAACCTGAATATCTGGACC  
ATCAAGGTTTCTGGTCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGTGCA  
GCCTCTGGACAACCCAAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA
```

Protein sequence: (SEQ ID NO: 217)

```
MKSGIATRRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPEVEKLAQAKETAGWKLVRQAFQGLHRKTSKNLNIWT  
IKVSGPRKTKELKAYLLQDPKLLFPVQPLDNPSLTVITDAEGGVE
```


FIGURE 31B

RS04

DNA sequence: (SEQ ID NO: 98)

```
ATAGACCAGTTGAGTGAGCAGGAGTCGGTGGAAAGTGGTCTGCTCAGCTTTCGATGTGGCGCGGTCTTGCTACTACGTCCA
CCGTCTTCGACCGCGCGGTGTGATGCTCGCCGCTGGCGCTACGCAGCCAAGTCAACCAGTTGTTTCAGCCAGAGTCGGG
GCTCGGCCGCGCAGCCGAGCATTCTGGGCATGCTGCGCGAAGAGGGCGTGACCATCGGCCGTTTCCGAGTGCCTCGGTTG
ATGCGTGAGCTGGGCTGGTCAGCAAGCAACCGGGCTCGCACGCCTACAAACAGGCCACGGTTGAGCGGCCGGATATCCC
GAATCGGCTGAACCGCAATTTCGCGACCGAGCATCCCATACAGGTGTGGTGTGGCGACATCACCTACGTCTGGGCGCAAG
GCCGTTGGCACTACCTGGCCGCGGTGCTGGATCTGCTGATCGGCTGGGCGTTCTCGGCCAAGCCGATGCCGAACCTGCTG
ATCAAGGCCCTGGACATGGCCTACGAACAGCGCGGCGAGGCCACAGCAGGTGCTGTTCCATTAGACCAGGGCAGCCAGTA
CGCCAGCCGCTGTTTCGGCAACGGCTCTGGCGCTATCGGATGCAGCAGAGCATGAGCCGTCGGGGGAATTGCTGGGATA
ACTCGCCGATGGAGCGCTGTTCCGAGTCTGAAGTCGGAGTGGGTCCCGTCAACGGGTTACCTGACGGCGCAGGAGGCC
CAACGGGACATCAGTCATTACTTGATGCACCGCTACAACCTGGATCAGGCCGATCAATTCAACGACGGGTTACCACCTGC
GGTGGCCGAAGAAAACTCAACCCACTGTCCGGGATGGGTTGA
```

Protein sequence: (SEQ ID NO: 218)

```
IDQLSEQESVEVVCSAFDVARSCYYVHRLRRRRVDARRVALRSQVNLFSQSRGSAGSRSILGMLREEGVITGRFRVRL
MRELGLVSKQPGSHAYKQATVERPDIPNRLNREFATEHPIQVWCGDITYVWAQGRWHYLAAVLDLLIGWAFSAKPDALV
IKALDMAYEQRGRPQQVLFHSDQGSQYASRLFRQLRWYRMQMSRRGNCWDNSPMERLFRSLKSEWVPSTGYLTAQEA
QRDISHYLMHRYNWIRPHQFNDGLPPAVAEKLNPLSGMG
```

RS05

DNA sequence: (SEQ ID NO: 282)

```
ATGAGCAAGCAACGACGTACGTTTTCCGCCGAGTTCAAACGAGAGGCCCGGCCCTGGTGTGGACCAAGGTACAGCCA
TATCGACGCCTGCCGTTTCGCTGGGGTGGTGGATTTCGGCCTTCGCGCGTTGGGTGAAGCAGCTCGAGGGCGGAGCGCCAGG
GTGTGACCCCGAAGAGCAAGCGCTTGACGCCTGAGCAGCAAAAGATCCAGGAGCTGGAAGCCCGGATCAACCGATTGGAG
CGGGAGAAAGCGATATTAAAAAGGCTACCGCTCTCTTGATGTCGGACGAACTCGATCGTACGCGCTGA
```

Protein sequence: (SEQ ID NO: 219)

```
MSKQRTFSAEFKREAAALVLDQGYSHIDACRSGLGVDSALRRWVKQLEAERQGVTPKSKALTPEQQKIQELERINRLE
REKAILKKATALMSDELDRTR
```

RS06

DNA sequence: (SEQ ID NO: 99)

```
ATGTTGTATTTTTCTTGCAGTATGAAGATGGGTGGTGGGTGGGATATAGGTACTTCTCTCTATTTTCTTTAATTGCTCT
CATCTATGGGTGTGTCGGTGGTGGAGGTGGATCGGATGAGATTGGGCAGCACTGCTTTGAGAGAGAGCAAAAGCTTTCCG
GAGTTAATGATAATGAAGAGGGGAGTGTGAGGTTGAATCGGCTGAACCTGCGATCCAATTGAAGGTCGTGTTCTTGAATCA
GAGAAGCTGATAAGAAAGCCGCCAATGAGCTGGGTATTCACTGA
```

Protein sequence: (SEQ ID NO: 220)

```
MLYFSCSMKMGWVGYRYFSLFSLIALIYGCVGGGGSDIYGQHFEREQKLSGVNDNEEGSVRLNRLNCDPIEGRVLES
EKLIRKPPNELGIH
```

RS07

DNA sequence: (SEQ ID NO: 100)

```
ATGAAAAATCACTTGTTATGTGCGGCTGTGCTTTTGGTGGCTAGCAATTCGCGTGTGCTGATGAGGGCTCAAATGATGG
AAGTGAGATATGTCGGGCGCAGGGTGGAGTTGAAATAACAAGTCTGGGGGAAGTCTCAAAGGTGTGGATGTTGAAGATG
TTGTAGTTTGTTCGATTCTTCCAAGTAATATGAAGTCGAGTCAAAGAGCGCTACACTCCCTCCTCTGCAAAGGATGATC
ATTTCCGGCAATGCCTTACCAGGAACGGTCACTGTTTCTGCCAGCGGAGATAGGAAATTTACAACATCTTGCCGGGCAAA
TCTTTATGCTCCACGTTATGCCAATTCTATCCAGACGGTGTTAGCAGGGGAACATCAGATCTACGATGTGTTGGTTACA
ATACACCCGGGAATTCATCTCAAGGTTGTAATGTGTCATGGGACGGCCCGACCGACATTCATTTGGGTGTTGAGCCATAT
GGCGGATCTGTTGTTGTAACATACAGTTGCAGTGCATTCAAAACAACGATTCAGTGATAATGAGCTACAGTTATCGTGA
TGGCGCGGCAGTGTATGGCGAGGTCAGAAATGTGTCAGGAATAATAAATGTGGTTTTGAACATAA
```

FIGURE 31C

Protein sequence: (SEQ ID NO: 221)

MKKSLVMSAVLLVASNFACADEGSNDGSEICRAQGGVEITSLGEVSKGVDVEDVVVCSILPSNMKSSQRAPTLPLQRFMI
ISAMPSPGTVTVSASGDRKFTTSCRANLYAPRYANFYPDGVSRGTSDLRCVGYNTPGNSSQGCNVSWDGPTDIQLGVEPY
GGSVVVNYSCTAFKTTIPVIMSYSYRDGRAVYGEVQNVSGIINVVLN

RS08

DNA sequence: (SEQ ID NO: 101)

ATGCTTATTAAATTCCTCGAATTATATTCTTGTTCCTATAGTTGGTTTGGCACAGCAGGCTGCTGCCTCCCCGCCCCGC
AGAGTCACACTCGGAACAATCTGAATCTTCGTGTATCGATGTCCAAGTCAATGGAGCACGTAGCCTGTCTTATAACTGCA
TGGCTCAGCAAATGACTCCACCCAAAGAGGATCCTCGGCGTCGGAACCTACCTTGAACCTCCACATTAGCGTCTGAACGC
GCCACTCGCCTGCCACCCACACAGACAGGACTTTTACCAGCCTTCATCAACGTGCCATATCGAACTCGAAAGACTAG

Protein sequence: (SEQ ID NO: 222)

MLIKILRIIFLLPIVGLAQQAASPPAESHSSEQSESSCIDVQVNGARSLSYNCMAQQMTPPKEDPRRRNPTLNSTLASER
ATRLPPTQTGLFTSLHQRAISNSKD

RS09

DNA sequence: (SEQ ID NO: 102)

GTGAGTAGTACTAAGAGTAAGCCGATAGCCAGGGGGCGTGGTGGCCCATTTGGGGAAGTGATGAAGAGGTGCGGGCTTGT
ACCGGTTTCGAGGAAGGAATAGACAGCAGACAGGATCGCTTGCGATGGGGCAGCAGGAAACCATCAGCCCGTCCGTATCCA
GAAGTGTGCTTGCAGCGTTAGGGGTGACTCCCTCATGCCCTAG

Protein sequence: (SEQ ID NO: 223)

VSSTKSKPIARGGGPFGEVMKRCGLVPVRGRNRQQTGSLAMGQQETISPSVSRTAACSVRGDSLMP

RS10

DNA sequence: (SEQ ID NO: 103)

ATGGAACGCTTGCTCGAGAGCATTTACATCAATGCCCGGCCGCGATGGAGTTGAGGCTTAGCCTCACCAGCTCCGGCCG
CAAGAGAATGGTAAAGATTGTGGATGGGAGGAGGTCGAGGTTCTGCCAGGTGAAGTGCAGGGCATCCTGGAGGCCCAAA
AGAGGGATGTTGGAATCTCGCCGACTTCTTAGCCAAGAGTCTCGTGGCGCGACGCTAG

Protein sequence: (SEQ ID NO: 224)

MERLLESIIYNARFAMELRLSLTSSGRKRMVKIVDGEVEVLPGEVQGILEAQKRDVGILADFLAKSLVARR

RS011

DNA sequence: (SEQ ID NO: 104)

ATGGAATGCCACGTTTCGTCCCGCCACGAGCAGAGATGCAGCAGCGATAAGCTGCGTAGTTATAGCCGCCCTGCGTGAGTC
AAATTCACAGGACTATCCGCCTGATGTGATCGCTCAGGTTGAGCAGAGCTTTTCTCCTGAAGCCATCACCACACAGCTTA
CGAAGCGTAGGGTCTTCGTAGCCTTATTGGGCGAAAACATTATTGGCACTGCCGGTCTCGACGGTGACGTCTGAGAAAT
GTTTTCGTTGACCCAGCTCACCAGAAAGGCGGTATCGGGCGGCATTGATGGATGTCATTATACAACTGCTGCCAGCGC
GGGAGTTGGAGCTGTACGTGTCCATCGTCGATTACAGCTGAAAGGTTTTATACCGCATTGGGTTATCAGAAAATCCGCG
ACGAGTTTCATGGGGCGGAGCGCACCATCGTTATGGAGAAGCGGCTGTAG

Protein sequence: (SEQ ID NO: 225)

MECHVRPATSRDAAAIISCVVIAALRESNSQDYPPDVIAQVEQSFSP EAITTQLTKRRVFVALLGENIIGTAGLDGDVVRS
VFVDPAHQKGGIGRHLMDVIHTTAASAGVGAVRVPSSITAERFYTALGYQKJRDEFHGAERTIVMEKRL

FIGURE 31D

RS12

DNA sequence: (SEQ ID NO: 105)

```
TTGTGGTTGACCTGCACGCCACAGCAGGATGTGCAGGCGGCGTTAGCTACAGCGTCGATACTCCTGGGCCAGTTCCACCA
GTGGGGCGTGCAGCTCGGTTCGCTACACTAGCCTCGACCCGCTTGAGGAAGTCGAGAAGAAGCCTTCTGCACTGCCGTCTC
CTGCTTGGAAAACGGATTCTACTAAGTTACAGCGTGGTACTGAAATCGGGGGGAGGTCAATCGACAAAGGTATCCCGACC
GCAGGTTTGTGGCCACGTGATGGTGGCCAAGTTTGCCGATCACTTGCCGCTGTACCGGCAGGAGAAAATCTTTGGCCG
CGCCGGGCTGGCAATTGCTCGCTCGACCTGGGCGAGTGGGTGGGACAAACCGGCGTTCGGCTTCAGCCACTGGTTCGATG
CACTGCGTGAAGCCGTCTGAACAGGGCGTGATCCACGCTGATGAAACACCGGTGCAAATGCTTGCGCCAGGCGAGAAG
AAAACCCACCGGGCCTATGTCTGGGCGTACAGCAGCAGCGCGTTTTCAGGGCTCAAAGCGGTGGTTTACGACTTCAGCCC
AAGCCGTGCTGGCGAATCGCGCAACTTCTGGGTGACTGGAACGGCAAGCTGGTCTGCGACGACTTCGCTGGCTACA
AAGCCGTTTTCGAACAAGGCATCACTGAAATCGGCTGCATGGCCCCACGCCCGGCGCAAGTTCTTTGATTTCAGCTGGCG
AACAAAGTCAGCTGGCTGAACAGGCCCTGCACTCGATCAGCGGCTTGACGAGGTGCAACGTCAGGCGCGGACATGAG
TGATGAAGAGCGCTGGCGAATACGACAAGAATTGGCGGTGCCGATCCTCAAAAACTGCATGACTGGATGTTGGCTCAGC
GAGACCTGGTGGCCAAATGGATCAGCCAGGCCAAAGCCCTCGATTACAGCCTGAAACGCTGGGTAGCGCTGACGCGCTAC
CTGGACGATGGGGCTGTGCCATCGATAACAATCAGTTCGAGAACCATAACGCCCATGGGCGCTCGGGCGTTTCAACTG
GCTGTTTCCGGGTCGCTGCGCAGTGGTAAACGGGCGGCTGCAATCATGAGCCTGATCTAG
```

Protein sequence: (SEQ ID NO: 226)

```
LWLTCTPQDVQAALATASILGQFHQLGVQLGRYTSLDPLEEVEKNASALPSPAWKTDSTKFSVVLKSGGRSIDKGIPT
AGLLAHVMVAKFADHILPLYRQEKIFRAGLAIAIRSTLAQWVGQTGVRLQPLVDALREAVLNQGVIIHADETPVQMLAPGEK
KTHRAYVWAYSTTFPSGLKAVVYDFSPSRAGEHARNFLGDWNGKLVCDDFAGYKAGFEQGITEIGCMAHARRKFFDLHVA
NKSQLAELQALHSISGLYEVEVERQARDMSDEERWRIRQELAVPILKKLHDWMLAQRDLPNGSATAKALDYSLRKWVALTRY
LDDGAVPIDNNQVENQIRPWALGRSNWLFAGSLRSGKRAAAIMSLI
```

RS13

DNA sequence: (SEQ ID NO: 106)

```
ATGGTGAGGCGGCGGAGGGTTCGCGGTGGCGCGCAATGCCTGAGCCTGTGAGCGCACCGAACCCAGGTCTTGTCGATGGA
TTTCGCTCTTCGACGCGCTCAGCACTGGGCGACGGATCAAATGCCTGACGGTGGTTCGATGACTTCACCAAGGTGTTCGGTCG
ACATCTTGGTGGAGTACGGTATCAGCGGTTTTCGTGTCAAGCGGGCGCTGGACGAGATGGCGCGGTTTTCGTGGCTACCCG
CAGGCGATCCGCACCGACCGAGGGCCCCGAGTTACCCGCAAGGCGCTTGATCAGTGGGCTGTGAGCGTGACATCAAGTT
GAAGCTGATTGAGCCTGGCCAGCCACGACAGAGCGCCTTCATCGAGTCATTCAACGGAAGTTCCGGGGCGAATGCCTCA
ATGAGCACTGCTCGCTGGTTCGAAGCCAGAATCCGTATCCGCGGTTCGGCGGATTACAACGAGCACCGACACACAGCGCC
ATTGGCAATCTCTCCCGCGCAGAGCTTGCTGCGAAGTGGCGAACCAACCAGCAGCAGCTGAAGCGGGAAAAGTTGATATC
AACCCCATAG
```

Protein sequence: (SEQ ID NO: 227)

```
MVRRRVAVARECLSLSSAPNQVLSMDFVFDALSTGRRIKCLTVVDDFTKVSVDILVEYGISGFRVTRALDEMARFRGYF
QAIRTDQGPFTGKALDQWACQRDIKLLIQPGQPTQSFIIESFNGKFRGECLNEHCSLVEARIRIAAWRDYNEHRPHSA
IGNLSPAELAAKWRNTNQQLKREKLSTP
```

RS14

DNA sequence: (SEQ ID NO: 107)

```
ATGCATATCCAATCGTTGGGGGCTACTGCCTCCTCGCTGAATCAGGAGCCTGTGCAAAACCCGTCGCAGGCAGCGCATAA
GTCCGCCAGCTTTCGCTCAGGAACCTTCAGGGCAAGGTCTCGGGGTTCCTTAAGAGCACGCCGGGAATACCTTCCGGGA
AGTTGCCGGAAGCGTTAGCGACGTGCGTTTCAGCAGTCCCCAAGGGCAAGGGGAGTCCCGTACTCTGACTGACTCGGCA
GGGCGCGGCAGATCACTGCGCCAGTTGAGAACGGAGTCACCGAGCTACAGCTCAGTCGGCCACCATTGACCACTCT
GGTCTCAAGCGGCGGTGGTGCCTAAGGTGCGGCATACCCGGGAGCAATGCTGGCGCTAGAAGAGAAAGGCATGCTCGATG
GCATCCGACGATGTCGGGTTTCGTCGCTGGCGGCATCACCGCGCCCTTTTGGCCTCAGGTATGAGCCCGGCGCGGTTT
AAGACCCCTTTCGACAAGATGGATCTTATTTTCGCTGCTCGACAGCTCGAACAGAAAGCTGAAGCTGTTCCAAACATTAG
CAGCGAGATCGGCGCATCGCTGAAAAAGGGCTTGGGCAACAAGATCGGCGGCTTCTCTGAGTTGCTGCTCAATGTACTCC
CAGCATAGATTCGCGGGCTGAGCCCCTAGAACGCCCTATTGCGCGACGAGACACGCAAGGCCGTGCTCGGACAGATCGCT
ACGCATCCAGAGGTTGCAGCGACGCCAGCCGTTGCCGCCATCGCCAGCAGATTGCACTCCGGCTCCGGAGTCACCTTTGG
CGATCTAGATCGGTTGAGTGCTTACATTCCCAGATTAAAGACGCTGAACATCACAGGTACGCCCATGTTTCGAGGGGCGTC
CGCAATTAGTGGTGTTCATGCCAGCCACACACCGGATCTGGAGGTGCGCCAGGCGGCACATATCTCCGGTTCTTCCCA
GGAGTGTTCCAGAGGTGACGTTGAGTGATCAGCCGTACCAGGCCGGCGTAGAGTGACAGAATTCCAGGATGGCGGGGT
GATGATTAACGTGCCGCTCCCTGAGATGATCGACAAGAATTTTGACAGCGGGCCACTGCGGCGCAACGACAACCTGATCC
TTGAGTTTCAGGGGCGAAGCTGGGAGGTAGCGGCCGACCGAGGTACTAGGGGCGGCGCGCTCAAGGGCTGGGTCTGCGGG
GTGCTGCCCTGCAAGCGCGCGAAATGCTGCAGCTCGAGGGCTGGAGGAATTGCGCGAGCAAAACGTTGTGGTTCGGTT
GAAGAGCGAGCGCGGTGATTTCACTGGCATGCTCGGTGGCACCTTGAACCTTACCATTGCCGACGAGATCAAGGCGCATC
TTCAGGAGCGCCTCCAGGAGCGAGTGGTGAACATCTGGAGAAACGCTTTCAGGCTTCAGAGCGTCATACCTTCGCTTCT
```

FIGURE 31E

CTCGACGAGGCGCTGCTGGCACTTGATGACAGTATGCTCACCAGTGTGCTCAACAGAACCCGGAGATCACAGACGGGGC
GGTGGCTTTTTCGCCAGAAGGCGCGGGATGCGTTCACCGAGCTGACTGTGCTATCGTTAGCGCCAATGGCTTGGCGGGTA
GGCTCAAGTTGGACGAGGCTATGCGCTCCGCTCTTCAGCGACTCGATGCGCTGGCAGATACTCCGGAACGCCTAGCATGG
TTGGCAGCTGAGTTGAACCATGCTGATAACGTTGATCATCAGCAGTTACTCGATGCCATGCGCGGGCAGACGGTGCAGTC
GCCGCTGCTCGCGCTGCGTTAGCAGAGGCGCAGCGCCGAAAGTGGCGGTTATTGCCGAGAACATTCGTAGGAAGTTA
TCTTCCCTCTCTGTATCGCCCTGGCCAGCCGATTCCAACGTAGCTCTGTTACGTGCGGCGGAGGAGCAGCTACGGCAT
GCCACAGTCCGGCGGAAATCAATCAAGCGCTGAACGATATCGTCGACAACCTACTCGGCACGAGGCTTCTGCGTTTCGG
CAAACCTTGAGTTCGACTACCGTTGAGATGGCTAAGGCTTGGCGGAATAAGGAGTTCACATGATT

Protein sequence (SEQ ID NO: 228)

MHIQSLGATASSLNQEPVETPSQAAHKASASLRQEPGQGLGVALKSTPGILSGKLPESVSDVRFSSPQGQGESRTLTDSA
GPRQITLRQFENGVTLEQLSRPPLTSLVLSGGGAKGAAYPGAMLALAEKGMLDGIRMSGSSAGGITAALLASGMSPAAF
KTLSDKMDLISLLDSSNNKLLKFQHI SSEIGASLKKGLGNKIGGFSELLNVLPRIDSRAPLERLLRDETRKAVLGQIA
THPEVARQPTVAAIASRLQSGSGVTFGDLRLSAYI PQIKTLNITGTAMFEGRPQLVFNASHTPDLEVAQAAHISGSFP
GVFQKVSLSDPYQAGVEWTFQDGGVMINVPVPEMIDKNFDSGPLRRNDNLI LEFEGEAGEVAPDRGTRGGALKGWVVG
VPALQAREMLQLEGLEELREQTVVPLKSERGDFSGMLGGTLNFTMPDEIKAHLQERLQERVGEHLEKRLQASERHTFAS
LDEALLALDDSMILTSAQQNPEITDGAVAFRQKARDAFTELTVAIVSANGLAGRLKLDEAMRSALQRLDALADTPERLAW
LAAELNHADNVDHQQLLDAMRGQTVQSPVLAAALAEARRKVAIVIAENIRKEVIFPSLYRPGQPDNSVALLRRAEEQLRH
ATSPAEINQALNDIVDNYSARGFLRFKPLSSTTVEMAKAWRNKEFT

RS15

DNA sequence: (SEQ ID NO: 108)

ATGATTGATACATGGCTGGCACAGTGGGGCTTGAGACTTCCCTCGAGCAACGATGCCACGTTGCGGCTGCAACCGGCAGA
GGGACCGGAACCTGGTTATGGAGCGCCTCGAGGGCGGTTGGCTTTTCGTCGTCGAGTTGGGACTTGTGCCTTCAGGGTTAC
CGCTGGGTGTGATCTTGCAATTGTTACAAGTGAACCTCTCCATTCTCATCCTTGGCACCGGTGAAACTTGGCGCGGACGAT
GCCGCTAGACTTGTGCTCTGGGCTGAGGCACGTGATGGCGTTGACGATGTGGATGCACTGAACCGCTTGACGATAGGCT
GCGGGAAGGACATTACGATTAGTCCATTGCTAGAGCCACGGGTGAGTTGGTTCCAGCTCAGATACAAACCAGCGCGT
TAGTGTTCGTTTGA

Protein sequence: (SEQ ID NO: 229)

MIDTWLAQWGLRPLSSNDATLRLQPAEGPELVMERLEGGWLFVVELGLVPSGLPLGVILQLLQVNSPFSSLPVKLAADD
AGRLVLWAEARDGVDDVDALNRLHDLRREGHSRLVPLLEPTGELVPAQIQTSALVFV

Big Island: Overall Nucleotide Homology

Total 84830bp

Bp #	Species, strain, gene name	Accession #	Evalue / %identity
23559-25465:	<i>X. axonopodis</i> pv. Citri strain 306	AE011864	83%

Only stretches within the 20437-25465bp are homologous to *X. axonopodis* pv. Citri strain 306.
 A total of 1060bp, not contiguous, from this region are homologous to *X. axonopodis*.

33872-38412:	<i>P. aeruginosa</i> , PA14, pvrR	AF482691	0.0; 99%
--------------	-----------------------------------	----------	----------

40989-46535:	<i>P. aeruginosa</i> , PA01, PA2128-2132	AE004640	80%
--------------	--	----------	-----

Only stretches within the 40989-46535bp region are homologous to PA01. A total of 2406bp, not contiguous, from this region are homologous to PA01.

48266-49533:	<i>P. putida</i> , plasmid pWWO	AJ344068	96%
--------------	---------------------------------	----------	-----

Only stretches within the 48266-49533bp are homologous to *P. putida*, plasmid pWWO. A total of 780bp, not contiguous, from this region are homologous to *P. putida*.

56824-58706:	<i>P. syringae</i> pv. <i>maculicola</i> , plasmid pFKN	AF359557	83%
--------------	---	----------	-----

Only stretches within the 56824-58706bp are homologous to *P. syringae* pv. *maculicola*, plasmid pFKN. A total of 1882bp, not contiguous, from this region are homologous to *P. syringae*.

64748-64942:	<i>P. aeruginosa</i> , PA103, exoU,	U97065	1E-85/96%
--------------	-------------------------------------	--------	-----------

82447-85179:	<i>P. aeruginosa</i> , PA01, PA0984-0985	AE004531	0.0/97%
--------------	--	----------	---------

85334-855542:			3E-80/94%
---------------	--	--	-----------

93200-93317:	<i>P. aeruginosa</i> , PA158	X73064	7E-50/98%
--------------	------------------------------	--------	-----------

108075-108610:	<i>P. aeruginosa</i> , SG17M, plasmid pKLC102	AF285416	0.0/91%
----------------	---	----------	---------

100119-101054:	<i>P. aeruginosa</i> , PA01, PA3849,	AE4802	0.0/98%
----------------	--------------------------------------	--------	---------

Figure 32

Small Island: Overall Nucleotide Homology

Total 10848bp

Bp #	(Species, strain, gene name	Accession #)	E value / %identity
877-1632:	<i>P. aeruginosa</i> , PA01, PA0977	AE004531	0.0/94%
877-2603:	<i>P. aeruginosa</i> , SG17M plasmid pKLC102	AF285416	0.0/92%
2978-6471:	<i>P. aeruginosa</i> , PA01, PA0978-81	AE004531	0.0/99%
7035-7999:	<i>P. syringae</i> pv. <i>maculicola</i> , plasmid-pFKN	AF359557	83%
Only stretches within the 7035-7999 bp are homologous to <i>P. syringae</i> pv. <i>maculicola</i> , plasmid pFKN.			
A total of 534bp, not contiguous, from this region are homologous to <i>P. syringae</i> .			
7999-8284	<i>P. aeruginosa</i> , PA103, <i>exoU</i> ,	AF27291	E-136/96
8000-8080:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-18/91%
8120-8259:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-24/85%
8272-8860:	<i>P. aeruginosa</i> , PA01 intragenic region	AE004531	E-176/88%
8470-11724	<i>P. aeruginosa</i> , PA103, <i>exoU</i> ,	AF27291	0.0/99%

FIGURE 34A

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (gene name / species strain)	E-value (%) identity	Proteins with lesser identity Cut off 30%	GenBank accession no.
RL001	-	878	1,202	pseudogene	56			Hypothetical protein	PA0977 / <i>P. aeruginosa</i> PAOI	3E-45 (91)		AA004366
RL002		11,69	2,452	427	60	cytoplasm	<i>gacA</i>	Integrase, <i>P. aeruginosa</i> SGT1M	Integrase, <i>P. aeruginosa</i> SGT1M	4.0 (91)	STY24608	AA02088
RL003	-	2,449	4,368	639	58	cytoplasm		Hypothetical protein	PSPT00831 / <i>P. syringae</i> p.v. <i>tomato</i> DC3000	E-180 (52)	XAC2196, XCC3121, STY4665	AA054371
RL004		4,903	5,486	177	59	cytoplasm	<i>gacA</i>	Deoxyxylidine deaminase	WK1566 / <i>Neisseria meningitidis</i> <i>Kandier</i> (LV19)	1E-13 (93)		AA02479
RL005	+	5,436	6,146	236	51	inner membrane		Hypothetical protein	No significant similarity			
RL006		61140	61682	179	45	cytoplasm	<i>gid</i>	Deoxyxylidine diphosphate deaminase	VN602456 / <i>Halobacterium</i> sp. NRC-1	1E-07 (92)		AA01883
RL007	-	6,887	7,402	171	48	outer membrane and periplasm		Hypothetical protein	No significant similarity			
RL008		1,829	9760	641	60	cytoplasm		Hypothetical protein	Protein fused from putative helicase (<i>Melanoarctium acetivagus</i> C2A) and hypothetical protein PA1935 (<i>P. aeruginosa</i> PAOI)	7E-09 (92) 8E-16 (92)		AA05388 / AA00532
RL009	-	9,757	12,180	807	60	inner membrane		Hypothetical protein	PA1939 / <i>P. aeruginosa</i> PAOI	2E-69 (30)		AA005327
RL010		12,553	12,660	100	54	inner membrane		Hypothetical protein	No significant similarity			
RL011	-	13,102	13,452	116	60	cytoplasm	<i>parE</i>	Plasmid stabilization protein, <i>Agrobacterium tumefaciens</i> C58	AGR_C2415p / <i>Agrobacterium tumefaciens</i> C58	8E-22 (45)		AAK87104
RL012		13,456	13,728	90	62	cytoplasm		Putative transcription regulator	AGR_C2415p / <i>Agrobacterium tumefaciens</i> C58	2E-18 (41)	SIF3093	AA057105
RL013	+	13,847	14,191	114	48	inner membrane		Hypothetical protein	No significant similarity			
RL014		14,217	15,728	503	58	inner membrane		Hypothetical protein	XAC2186 / <i>P. aeruginosa</i> p.v. <i>tomato</i> DC3000	1E-75 (52)	XGG3117, STY4579	AA057059
RL015	-	15,725	16,066	113	62	inner membrane		Hypothetical protein	No significant similarity			
RL016		16,066	17,448	460	64	outer membrane and periplasm		Hypothetical protein	PSH00848 / <i>P. syringae</i> p.v. <i>tomato</i> DC3000	8E-18 (57)	XGG3116, SIF4377	AA057383

FIGURE 34B

RL017	-	17,466	18,404	312	65	outer membrane and periplasm		Hypothetical protein	PSPT00849 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	E-125 (72)	XAC2283, STY4576	AAO54384
RL018	-	18,404	18,836	141	62	outer membrane and periplasm		Hypothetical protein	SG524 / <i>P. aeruginosa</i> SG11M	E-152 (49)	STY4576	AAO54385
RL019	+	19,044	19,262	72	55	cytoplasm		Hypothetical protein	No significant similarity		XAC2275	
RL020	-	19,259	19,918	219	62	outer membrane and periplasm	ABC	Putative protein disulfide isomerase	PA09821 / <i>P. aeruginosa</i> PAO1	E-192 (93)		AAO54391
RL021	-	19,915	20,199	94	57	cytoplasm		Hypothetical protein	PSPT00858 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	2E-23 (55)	STY4573, STY4572	AAO54393
RL022	-	20,199	20,428	980	64	cytoplasm		Hypothetical protein	PSPT00859 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	0.0 (64)		AAO54394
RL023	-	23,138	23,581	147	64	inner membrane		Hypothetical protein	PSPT00860 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	2E-48 (74)		AAO54395
RL024	-	23,559	25,064	501	68	outer membrane and periplasm		Hypothetical protein	PSPT00861 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	E-148 (54)	STY4570, STY4569	AAO54396
RL025	-	25,048	25,932	294	66	outer membrane and periplasm		Hypothetical protein	PSPT00862 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	1E-90 (62)	XAC2272, STY4569	AAO54397
RL026	-	25,929	26,588	215	60	inner membrane		Hypothetical protein	PSPT00863 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	E-171 (59)	STY4568	AAO54398
RL027	-	26,585	26,971	128	65	inner membrane		Hypothetical protein	XAC2271 / <i>X. axonopodis</i> pv. <i>citr</i> 306	4E-29 (50)		AAM37124
RL028	-	26,982	27,588	118	60	inner membrane		Hypothetical protein	ORF116 / <i>P. putida</i> (Glasgow BMMO)	6E-23 (50)	XAC2270, STY4567	GAO54404
RL029	-	27,356	27,595	79	63	inner membrane		Hypothetical protein	CS4 / <i>P. aeruginosa</i> C	1E-09 (38)	STY4565	AAN62148
RL030	-	27,592	27,951	109	66	inner membrane		Putative type III effector protein	PSPT00869 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	E-125 (63)	XAC2270, STY4567	AAO54404
RL031	-	28,024	28,329	101	57	cytoplasm		Hypothetical protein	No significant similarity		PA4736, PA4737	
RL032	-	28,502	28,816	105	49	outer membrane and periplasm		Hypothetical protein	PA0718 / <i>P. aeruginosa</i> PAO1	9E-16 (43)		AAO54402
RL033	-	28,810	29,967	385	47	cytoplasm		Hypothetical protein	No significant similarity			
RL034	-	30,098	31,579	493	59	inner membrane		Putative DNA helixase	PSPT00879 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	0.0 (65)		AAO54413

FIGURE 34C

RL035	-	31,590	32,240	216	62	inner membrane		Hypothetical protein	PSP00880 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	2E-74 (63)	XAC2260, STY4563	AA054414
RL036		32,592	33,567	911	60	inner membrane		sensor of two-component regulatory system	ORF1/Pseudomonas PA14 / <i>ResG</i> / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhimurium	0.0 (91) / 1E-598 (92)	VES3	AA055582 / GAD04505
RL037	+	35,364	36,563	399	59	cytoplasm	<i>pvrR</i>	component regulatory system, adhesion and antibiotic resistance	<i>PvrR</i> / <i>P. aeruginosa</i> PA14 / <i>VieA</i> / <i>X. campestris</i> pv. <i>campestris</i> ATCC 33913	0.0 (100) / 2E-55 (34)	<i>pvrR</i>	AA015533 / AA041975
RL038		36,641	39,898	1086	63	inner membrane	<i>rcsG</i>	sensor of two-component regulatory system	STY0227 / <i>S. typhimurium</i> DT12	1E-85 (92)		AA021172
RL039	+	39,898	40,593	231	59	cytoplasm	<i>rcsB</i>	regulator of two-component regulatory system	Z3476 / <i>E. coli</i> O157:H7 EDL933	8E-33 (39)		AA057352
RL040		40,697	41,553	228	67	outer membrane and periplasm	<i>cupD5</i>	Probable pili assembly chaperone / adhesion and protein secretion	PA2182 / <i>P. aeruginosa</i> PAO1	4E-63 (62)		AA055520
RL041	-	41,343	42,689	448	63	outer membrane and periplasm	<i>cupD4</i>	adhesion and protein secretion	PA2131 / <i>P. aeruginosa</i> PAO1	0.0 (70)	STY0370	AA055519
RL042		42,686	45,298	1870	66	outer membrane and periplasm	<i>cupD3</i>	Probable integral biogenesis / adhesion and protein secretion	PA2130 / <i>P. aeruginosa</i> PAO1	0.0 (67)	STY0371	AA055518
RL043	-	45,282	46,028	248	65	outer membrane and periplasm	<i>cupD2</i>	Probable pili assembly chaperone / adhesion and protein secretion	PA2129 / <i>P. aeruginosa</i> PAO1	1E-83 (64)	STY0372	AA055517
RL044		46,167	46,665	1183	65	outer membrane and periplasm	<i>cupD1</i>	Probable integral precursor / adhesion and protein secretion	PA2128 / <i>P. aeruginosa</i> PAO1	6E-66 (72)	STY0373	AA055516
RL045	-	46,836	47,101	pseudogene	55			Recombination	Transposase / <i>E. coli</i> (plasmid p1658/97)	2E-36 (82)		AA049572
RL046		47,103	47,819	241	61	inner membrane		Hypothetical protein	PSP00880 / <i>P. syringae</i> pv. <i>tomato</i> DC3000	1E-63 (59)	XAC0260, STY4563	AA055414
RL047	-	47,849	50,083	744	65	inner membrane		Hypothetical protein	SG68 / <i>P. aeruginosa</i> SG17M	0.0 (66)	XAC2259, STY4562	AA062290
RL048		50,087	50,914	65	59	cytoplasm		Hypothetical protein	Non-pigment similarity			

FIGURE 34D

RL049	-	50,353	50,853	166	66	outer membrane and periplasm	Hypothetical protein	SG69 / <i>P. aeruginosa</i> SG17M	7E-34 (51)	STY4560	AAN62291
RL050		50,350	51,411	159	64	outer membrane and periplasm	Hypothetical protein	SG70 / <i>P. aeruginosa</i> SG17M	5E-58 (46)	STY4559	AAN62292
RL051	-	51,416	52,171	251	65	outer membrane and periplasm	Hypothetical protein	C68 / <i>P. aeruginosa</i> C	5E-50 (52)	STY4558	AAN62162
RL052		52,182	52,871	221	64	inner membrane	Hypothetical protein	C69 / <i>P. aeruginosa</i> C	5E-05 (25)		AAN62163
RL053	-	53,019	53,235	pseudogene	61		Recombination protein	Transposase / <i>P. putida</i> (plasmid pWW53)	7E-09 (54)		BAB59051
RL054		53,482	54,515	241	47	cytoplasm	Hypothetical protein	PA2225 / <i>P. aeruginosa</i> PAO1	2E-50 (45)		AAG05601
RL055	+	54,524	55,189	221	50	cytoplasm	Hypothetical protein	PA2222 / <i>P. aeruginosa</i> PAO1	2E-56 (48)		AAG05610
RL056		55,272	55,946	221	49	cytoplasm	Hypothetical protein	PA2224 / <i>P. aeruginosa</i> PAO1	2E-06 (23)		AAG05612
RL057	+	56,030	56,275	81	59	cytoplasm	Hypothetical protein	No significant similarity			
RL058		56,388	56,707	119	48	cytoplasm	Hypothetical protein	QAC3557 / <i>Glaucidium</i> <i>acrobaticum</i>	2E-13 (34)	XGC0096	AAG05608
RL059	-	56,768	58,303	511	59	cytoplasm	Transposase	PP4439 / ISP-pu14 ORF3 / <i>P. putida</i> K12440	0.0 (90)		AAN70015
RL060		58,667	58,702	111	61	cytoplasm	Transposase	PP4438 / ISP-pu14 ORF2 / <i>P. putida</i> K12440	2E-59 (08)		AAN70014
RL061	-	58,753	59,019	88	57	cytoplasm	Transposase	PP4437 / ISP-pu14 ORF1 / <i>P. putida</i> K12440	7E-39 (88)		AAN70013
RL062		59,776	60,802	408	46	cytoplasm	Hypothetical protein	No significant similarity			
RL063	-	60,826	63,075	749	63	inner membrane	Plasmid-related protein	XAC2243 / <i>X. axonopodis</i> pV- <i>citr</i> 306	0.0 (69)		AAM37096
RL064		63,180	64,651	481	64	cytoplasm	Plasmid-related protein	XAC2242 / <i>X. axonopodis</i> pV- <i>citr</i> 306	3E-183 (65)		AAM37095
RL065	-	64,661	65,266	201	63	cytoplasm	Hypothetical protein	XAC2241 / <i>X. axonopodis</i> pV- <i>citr</i> 306	1E-54 (56)		AAM37094
RL066		65,588	66,612	84	60	cytoplasm	Hypothetical protein	XE1757 / <i>P. cystidinosus</i> 9a5c	9E-09 (40)	XAC2240	AAM84566
RL067	-	65,680	66,042	120	58	cytoplasm	Hypothetical protein	CT7 / <i>P. aeruginosa</i> C	8E-23 (42)	XAC2239	AAN62171
RL068		66,112	66,587	91	63	cytoplasm	Hypothetical protein	No significant similarity			
RL069	-	66,384	67,073	229	60	inner membrane	Hypothetical protein	XE1760 / <i>X. fastidiosa</i> 9a5c	3E-50 (53)	XAC2237	AAM84569

FIGURE 34E

RL070	-	67,077	67,427	116	62	cytoplasm		Hypothetical protein	No significant similarity			
RL071	-	67,663	68,370	235	53	cytoplasm		Hypothetical protein	SG91 / <i>P. aeruginosa</i> SG17M	6E-35 (40)		AAN62312
RL072		68,855	69,091	71	52	cytoplasm		Hypothetical protein	STY4535 / <i>S. enterica</i> subsp. <i>enterica</i> ser. <i>dipteric</i> (G118)	7E-11 (64)		GAD09214
RL073	+	69,111	69,377	88	57	cytoplasm		Hypothetical protein	No significant similarity			
RL074		69,426	69,965	179	50	cytoplasm		Hypothetical protein	XEFY614 / <i>S. enterica</i> subsp. <i>enterica</i> ser. <i>dipteric</i> (G118)	1E-28 (62)		AAN64570
RL075	-	70,626	71,192	188	58	cytoplasm		Hypothetical protein	No significant similarity			
RL076		71,191	71,835	214	60	cytoplasm		Hypothetical protein	No significant similarity			
RL077	-	72,107	72,544	145	68	outer membrane and periplasm	<i>pilM2</i>	Type IV B pilus / adhesion and protein secretion	<i>PilM</i> / <i>S. typhimurium</i> (plasmid R64)	5E-04 (21)	PA4199, STY4540	BAB91693
RL078		72,553	73,901	442	63	outer membrane and periplasm	<i>pilI2</i>	Type IV B pilus / adhesion and protein secretion	STY4550, XAC22151	3E-65 (68)		CP642
RL079	-	73,906	74,847	313	65	inner membrane	<i>pilI7</i>	Type IV B pilus / putative peptidase / adhesion and protein secretion	BfpF / <i>E. coli</i> (plasmid PB171)	8E-25 (27)	XAC2923, XCC2754, PA0396, PA0395	BAA84845
RL080		74,841	75,374	176	60	outer membrane and periplasm	<i>pilS2</i>	Type IV B pilus / adhesion and protein secretion	<i>PilS</i> / <i>Stigella</i> ser. <i>enterica</i> (plasmid G01bP9)	5E-15 (29)	STY4547	BAA95180
RL081	-	75,396	76,475	359	55	inner membrane	<i>pilR2</i>	Type IV B pilus / adhesion and protein secretion	<i>PilR</i> / <i>S. typhimurium</i> (plasmid R64)	4E-51 (33)	BfpE, STY4546, XAC0697, XCC3423 (Type II), PA2676, PA3102, STY0164	BAB91688
RL082		76,475	76,905	526	63	inner membrane	<i>pilQ2</i>	ATPase / Type IV B pilus / adhesion and protein secretion	<i>PilQ</i> / <i>S. typhimurium</i> (plasmid R64)		STY4545, BfpD, BfpE (Type II)	BAB91689
RL083	-	78,064	78,597	177	69	outer membrane and periplasm	<i>pilP2</i>	Type IV B pilus / adhesion and protein secretion	<i>PilP</i> / <i>S. typhimurium</i> (plasmid R64)	1E-06 (34)	STY4544	BAB91690
RL084		78,587	79,912	141	64	inner membrane	<i>pilQ2</i>	Type IV B pilus / adhesion and protein secretion	<i>PilQ</i> / <i>S. enterica</i> subsp. <i>enterica</i> ser. <i>dipteric</i> (G118)	3E-15 (21)	STY4543	AAN81211

FIGURE 34F

RL085	-	79,916	81,538	540	63	outer membrane	<i>pilZ</i>	Secretin / Type IV B pilus / adhesion and protein secretion	<i>PilN / E. coli K-12 (plasmid R721)</i>	6E-84 (35)	bfpB, PA1382 (Type II), XFI527, peD (general secretion pathway protein)	BAB12647
RL086		81,627	82,751	57	66	outer membrane and periplasm	<i>pilZ</i>	Type IV B pilus / adhesion and protein secretion	STY4539/ <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhimurium	1E-48 (34)		CAD093117
RL087	+	83,023	83,349	108	40	inner membrane		Colicin immunity protein	PA0984 / <i>P. aeruginosa</i> PAO1	4E-55 (94)		AAC04373
RL088		83,774	84,870	495	47	inner membrane		Colicin immunity protein	PA0985 / <i>P. aeruginosa</i> PAO1	0.0 (97)		AAG04374
RL089	-	85,558	87,531	657	64	inner membrane		DNA Helicase	<i>discoidium</i> AX14	4E-28 (24)	PA0799	AAO51158
RL090		87,528	89,417	629	60	cytoplasm		Hypothetical protein	No significant similarity			
RL091	-	89,551	90,021	156	44	cytoplasm		Similar to luminal binding protein	Ricr142 / <i>Rhizobium rhizogenes</i> MAF03-01724	5E-29 (41)		BAB16261
RL092		90,095	92,020	641	62	cytoplasm	<i>topA</i>	Topoisomerase I	XE30003 / <i>S. flexneriae</i> 9556 (plasmid pXB51)	1E-151 (43)	PA9011, XFI1726, XFI0920, STY41298, XG03755, STY19366	AAT85572
RL093	+	92,340	92,573	77	58	cytoplasm		Hypothetical protein	No significant similarity			
RL094		92,712	92,957	81	59	cytoplasm		Hypothetical protein	No significant similarity			
RL095	-	93,421	93,891	156	61	cytoplasm	<i>ssb</i>	Single-stranded DNA binding protein	CI02 / <i>P. aeruginosa</i> C	6E-32 (53)	XAC211, PA4232	AAN62318
RL096		93,905	94,458	177	60	cytoplasm		Hypothetical protein	XFI179 / <i>K. gasindosa</i> 9556	8E-39 (53)	XAC2210	AAN84587
RL097	-	94,444	95,172	242	63	cytoplasm		Hypothetical protein	ORF6 / <i>Pseudomonas</i> sp. B13	7E-66 (52)	XAC2209, STY4529	CAD60670
RL098		95,442	95,681	79	46	inner membrane		Hypothetical protein	No significant similarity			
RL099	-	95,672	95,947	91	60	cytoplasm		Hypothetical protein	No significant similarity			
RL100		95,944	97,269	441	60	cytoplasm		Hypothetical protein	ORF5 / <i>Pseudomonas</i> sp. B13	1E-68 (39)	XAC2208, STY4528	CAD60669
RL101	-	97,266	98,033	255	58	cytoplasm		Hypothetical protein	ORF4 / <i>Pseudomonas</i> sp. B13	2E-33 (40)	XFI782, STY4526	CAD60668
RL102		98,061	99,800	579	58	cytoplasm		Hypothetical protein	SG102 / <i>P. aeruginosa</i> SG17M	1E-116 (49)	XFI783, XFI784, STY4525	AAN62323
RL103	-	99,797	100,051	84	60	cytoplasm		Hypothetical protein	No significant similarity			
RL104		100,048	101,064	538	64	cytoplasm		Hypothetical protein	PA3849 / <i>P. aeruginosa</i> PAO1	0.0 (96)	STY2463, STY4607, STY46226	AAG07256
RL105	-	101,064	101,297	77	65	cytoplasm		Hypothetical protein	No significant similarity			

FIGURE 34G

RL106	-	101,290	101,724	104	62	cytoplasm	Hypothetical protein	No significant similarity			
RL107	-	101,777	102,034	85	60	cytoplasm	Hypothetical protein	No significant similarity			
RL108	-	101,031	102,558	135	62	cytoplasm	Hypothetical protein	No significant similarity			
RL109	-	102,745	104,107	pseudogene	61		DNA replication and recombination	Replicative DNA helicase DnaB / <i>Pseudomonas</i> sp. SLT2001 (plasmid pOBR55)	E-130 (60)	PA4931, STY4442, XAC1477, XF0361, XCC1434	CAD13464
RL110	-	104,282	104,989	235	60	cytoplasm	Plaintive protein	DNA/RNA helicase DnaB / <i>Pseudomonas</i> sp. SLT2001 (plasmid pOBR55)	E-130 (60)	PA4931, STY4442, XAC1477, XF0361, XCC1434	CAD13464
RL111	-	104,986	105,687	233	64	cytoplasm	Hypothetical protein	STY1595 / <i>S. enterica</i> subsp. <i>enterica</i> ser. Typhi CT18	E-16 (32)		CAD01841
RL112	-	105,687	106,573	228	64	inner membrane	Hypothetical protein	No significant similarity			
RL113	-	106,516	107,013	165	60	outer membrane and periplasm	Hypothetical protein	PA2226 / <i>P. aeruginosa</i> PAOI	2E-32 (47)		AAG05614
RL114	-	107,010	107,690	226	58	cytoplasm	Hypothetical protein	QRE502 / <i>P. putida</i> plasmid pDITG1	2E-28 (93)		AF471807
RL115	-	107,687	108,616	309	60	inner membrane	Chromosome partitioning	Soj / <i>P. aeruginosa</i> SGI7M	1E-84 (90)	XF1785, XAC2205, STY4521	AAG02083
		108,700	108,757		47			<i>ankR</i> / <i>P. aeruginosa</i> PAOI			

Alignment: No_2 - emb|AL039136|HSM003612

Homo sapiens mRNA; EST DKFZp566K094_r1 (from clone DKFZp566)

Q:	18	DQTCNLSQNPPHHLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH	SEQ ID NO: 230
		DQTCNLSQNPPHHLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH	SEQ ID NO: 231
H:	29	DQTCNLSQNPPHHLLRLLDHWGDPAGCWSLGQTYSGHLYLPYCRELHKCSLCAHRNWH	SEQ ID NO: 232
		HYCCLWPVWMLCYMSW	93
		HYCCLWPVWMLCYMSW	
		HYCCLWPVWMLCYMSW	256

Figure 35

Alignment: No_8 - embi|M79137|HSXT01285

EST01285 Subtracted Hippocampus, Stratagene (cat. #936205) H

Q: 18 QVQHPPFLCLLDQHQQECIPPCLPDHLQDPQHPFLLPDHHVPHLVVLIQPQLCRALAP SEQ ID NO: 233
75 QVQHP .CLLDQHQQECIPPCLPDHLQDPQHPFLLPDHHVP.LVVLIQPQLCRALAP SEQ ID NO: 234
H: 43 QVQHPXXCLLDQHQQECIPPCLPDHLQDPQHPFLLPDHHVFXLVVLIQPQLCRALAP SEQ ID NO: 235
216

Figure 36

Alignment: No47 - swissnew|P35555|FBN1_HUMAN

FIBRILLIN 1 PRECURSOR.//:swiss|P35555|FBN1_HUMAN FIBRILLIN 1
 PRECURSOR.//:trembl|L13923|HSFIBRLN_1 product: "fibrillin"; Homo sapiens fibrillin
 mRNA, complete cds. //:gp|L13923|306746 product: "fibrillin"; Homo sapiens fibrillin mRNA,
 complete cds.

Q: 18 CCGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:236
 --- CCGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:237
 H: 2617 CCGASCHNTLGSYKCMCPAGFQYEQFSGGCQDINECGSAQAPCSYGCSNTEGGYLCGCPP SEQ ID NO:238

GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKGRGRKRRSTNET SEQ ID NO:236
 GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKGRGRKRRSTNET SEQ ID NO:237
 GYFRIGQGHCVSGMGMGRGNPEPPVSGEMDDNSLSPEACYECKINGYPKGRGRKRRSTNET SEQ ID NO:238

DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV-NKVRIL 178 SEQ ID NO:236
 DASNIEDQSETEANVSLASWDVEKTAIFAFNISHV NKVRIL SEQ ID NO:237
 DASNIEDQSETEANVSLASWDVEKTAIFAFNISHVSNKVRIL 2778 SEQ ID NO:238

Figure 37

Alignment: N056 - trembl|AF088916|AF088916_1

gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

//:trembl|AF088916|AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR. //:gp|AF088916|5353510 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.

//:gpnew|AF162780|6693840 gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

Q: 7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO: 239
 DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA:::.GY. E LE SEQ ID NO: 240
 H: 896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLE- SEQ ID NO: 241
 HRPPGALHTCGGP-GAFHLIVHLKAGDAV 94 SEQ ID NO: 239
 ::P .. :.. G. G.F.LI: L:AGD.V SEQ ID NO: 240
 NKPVAESQSPGTLGVFSLILPLQAGDTV 983 SEQ ID NO: 241

gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

//:trembl|AF088916|AF088916_1 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR. //:gp|AF088916|5353510 product: "emilin precursor"; Homo sapiens emilin precursor, mRNA, complete cds and 3' UTR.

//:gpnew|AF162780|6693840 gene: "EMI"; product: "elastin microfibril interfase located protein"; Homo sapiens elastin microfibril interfase located protein (EMI) gene, complete cds.

Q: 7 DGDVYNPSTGVFTAPYDGRYLITATLTPERDAYVEAVLSVSNASVAQLHTAGYRREFLEY SEQ ID NO: 242
 DG..Y:P.TGVFTAP. GRYL::A.LT .R.. VEAVLS SN..VA:::.GY. E LE SEQ ID NO: 243
 H: 896 DGGYYDPETGVFTAPLAGRYLLSAVLTGHRHEKVEAVLSRSNQGVARVDSGGYEPEGLEN SEQ ID NO: 244
 HRPPGALHTCGGPGAFHLIVHLKAGDAV 94 SEQ ID NO: 242
 . . : : G G.F.LI: L:AGD.V SEQ ID NO: 243
 KPVAESQSPGTLGVFSLILPLQAGDTV 983 SEQ ID NO: 244

Figure 38

Alignment: No59 - pironly|A35763|A35763

unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha) mRNA, complete cds.
 //:pironly|A35763|A35763 collagen alpha 2 chain - sea urchin (Paracentrotus lividus)
 (fragment)//:gp|J05422|159962 unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha)
 mRNA, complete cds.

Q:	92	GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL	142	SEQ ID NO: 245
		GE G.SG...P QG:RGI G: G.... GPR P . GGG S G.. GL		SEQ ID NO: 246
H:	718	GEPGPSENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGGSRGERGGL	768	SEQ ID NO: 247

unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha) mRNA, complete cds.
 //:pironly|A35763|A35763 collagen alpha 2 chain - sea urchin (Paracentrotus lividus)
 (fragment)//:gp|J05422|159962 unnamed ORF; P.lividus 2-alpha collagen (COLL2-alpha)
 mRNA, complete cds.

Q:	92	GENGSSGSQAPLQGLRGIFGLWGRRSRARFCGPR-PVARLGGGTSAGRELGL	142	SEQ ID NO: 248
		GE G.SG...P QG:RGI G: G.... GPR P . GGG S.G.. GL		SEQ ID NO: 249
H:	718	GEPGPSENGP-QGVRGIPGVVGENGKTGRGGPRGPPGLRGGGGSRGERGGL	768	SEQ ID NO: 250

Figure 39

Alignment: No60/63 - swiss|P20062|TCO2_HUMAN

TRANSCOBALAMIN II PRECURSOR.//:treml|M60396|HSTCII_1 gene: "TCN2"; product:
 "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
 //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II
 (TCII) mRNA, complete cds.

Q:	8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO: 251
		VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO: 252
H:	183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMAIRTVREEILKAQTPEGHFGN	SEQ ID NO: 253
		VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO: 251
		VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO: 252
		VYSTPLALQFLMTSPMPGAELGTACLKARVALLASLDGAFQNALMISQLLPVLNHKTYI	SEQ ID NO: 253
		DLIFPDCLAPRVMLEPAA	145 SEQ ID NO: 251
		DLIFPDCLAPRVMLEPAA	SEQ ID NO: 252
		DLIFPDCLAPRVMLEPAA	320 SEQ ID NO: 253

TRANSCOBALAMIN II PRECURSOR.//:treml|M60396|HSTCII_1 gene: "TCN2"; product:
 "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
 //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II
 (TCII) mRNA, complete cds.

Q:	8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	47 SEQ ID NO: 254
		VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	SEQ ID NO: 255
H:	183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	222 SEQ ID NO: 256

TRANSCOBALAMIN II PRECURSOR.//:treml|M60396|HSTCII_1 gene: "TCN2"; product:
 "transcobalamin II"; Human transcobalamin II (TCII) mRNA, complete cds.
 //:gp|M60396|339196 gene: "TCN2"; product: "transcobalamin II"; Human transcobalamin II
 (TCII) mRNA, complete cds.

Q:	8	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	47 SEQ ID NO: 257
		VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	SEQ ID NO: 258
H:	183	VEPFHQGHHSVDTAAMAGLAFTCLKRSNFPGRQRITMA	222 SEQ ID NO: 259

Figure 40

Alignment: No65 - swissnew|P23142|FBL1_HUMAN

FIBULIN-1 PRECURSOR.//:swiss|P37888|FBLD_HUMAN FIBULIN-1, ISOFORM D
 PRECURSOR.//:treml|U01244|HS2444_1 product: "fibulin-1D"; Homo sapiens fibulin-1D
 mRNA, complete cds. //:gp|U01244|1621019 product: "fibulin-1D"; Homo sapiens fibulin-1D
 mRNA, complete cds.

Q:	18	RNCQDIDECVTGIHNCSINETCFNIOGGFRCLAFECPENYRRAATLQOEKTDTVRCIKS	SEQ ID NO: 260
		RNCQDIDECVTGIHNCSINETCFNIOG.FRCLAFECPENYRRAATLQOEKTDTVRCIKS	SEQ ID NO: 261
H:	521	RNCQDIDECVTGIHNCSINETCFNIOGAFRCLAFECPENYRRAATLQOEKTDTVRCIKS	SEQ ID NO: 262
		CRPNDVTCVFDPVHTISHTVISLPTTFREFTRP EEIIFLRAITPPHPASQANIIFDITEGN	SEQ ID NO: 260
		CRPNDVTCVFDPVHTISHTVISLPTTFREFTRP EEIIFLRAITPPHPASQANIIFDITEGN	SEQ ID NO: 261
		CRPNDVTCVFDPVHTISHTVISLPTTFREFTRP EEIIFLRAITPPHPASQANIIFDITEGN	SEQ ID NO: 262
		LRDSFDIIKRYMDGMTVGIRR	158 SEQ ID NO: 260
		LRDSFDIIKRYMDGMTVG: R	SEQ ID NO: 261
		LRDSFDIIKRYMDGMTVGVR	661 SEQ ID NO: 262

Figure 41

Alignment: 80 - trembl|AF045447|AF045447_1

gene: "DPC4"; product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in
 pancreatic carcinoma (DPC4) gene, exon 11 partial sequence and complete cds.
 //:trembl|U44378|HS443781_1 gene: "DPC4"; product: "Dpc4"; Human homozygous deletion
 target in pancreatic carcinoma (DPC4) mRNA, complete cds. //:pironly|S71811|S71811
 probable transcription regulator MAD-4 - human//:gp|AF045447|2865657 gene: "DPC4";
 product: "deleted in pancreatic carcinoma"; Homo sapiens deleted in pancreatic carcinoma
 (DPC4) gene, exon 11 partial sequence and complete cds. //:gp|U44378|1163234 gene:
 "DPC4"; product: "Dpc4"; Human homozygous deletion target in pancreatic carcinoma
 (DPC4) mRNA, complete cds.

Q:	6	PGSRIRGRVDTLQXNAPXXMMVKDEYVHDFEGQPXLXTEGHXIQTIQHPPXNRAXTETYY	SEQ ID NO: 263
		PG :.G TLQ.NAP..MMVKDEYVHDFEGQP.L.TEGH.IQTIQHPP.NRA.TETY.	SEQ ID NO: 264
H:	139	PGIDLGL--TLQSNAPSSMMVKDEYVHDFEGQPSLSTEGHSIQTIQHPPSNRASTETYS	SEQ ID NO: 265
		TPALLAPXEXNATXTANFPNIPVAXTXQPAXILGGXHXEGLLQIAXGPQPGQQQNGFTGQ	SEQ ID NO: 263
		TPALLAP.E.NAT.TANFPNIPVA.T.QPA.ILGG.H.EGLLQIA.GPQPGQQQNGFTGQ	SEQ ID NO: 264
		TPALLAPSESATSTANFPNIPVASTSQPASILGGSHSEGLLQIASGPQPGQQQNGFTGQ	SEQ ID NO: 265
		PATYHHNXTTTWTGXTAPYTPNLPHHQKG	155 SEQ ID NO: 263
		PATYHHN.TTTWTG.RTAPYTPNLPHHQ.G	SEQ ID NO: 264
		PATYHHNSTTTWTGSRTPYTPNLPHHQNG	286 SEQ ID NO: 265

Figure 42

Alignment: No86 - trembl|D32210|D32210_1

gene: "Notch2"; product: "cell surface protein"; Mus musculus (Notch2) mRNA, complete cds.
 //:gp|D32210|2373395 gene: "Notch2"; product: "cell surface protein"; Mus musculus
 (Notch2) mRNA, complete cds.

Q:	81	MPALRPALLWALLALWLCCATPAHALQCRDGYEPCVNEGMCVITYHNGTGYCKCP-GFLGE	SEQ ID NO: 266
		MP LRPA.L ALL LWLC A PAHALQCR.G.EPCVNEG.CVITYHNGTG:C:CP GFLGE	SEQ ID NO: 267
H:	1	MPDLRPAALRALLWLWLCGAGPAHALQCRGGQEPVNEGTCVITYHNGTGFCRCPEGFLGE	SEQ ID NO: 268
		YCQHR-PCEKNRCGDPSTC	157 SEQ ID NO: 266
		YCQHR PCEKNRC : .TC	SEQ ID NO: 267
		YCQHRDPCEKNRCQGGTC	79 SEQ ID NO: 268

Figure 43

AATTCGCGGCCGCGTCGACGACCAACCTGTGATAACCTGCTCATAAACCTCCTCATCATCTACTCCTTCGTCT 75

SEQ ID NO: 109

adaptor

. I R G R V D D O T C D N L S O N P P H H L L L R L

SEQ ID NO: 269

TCTGGATCACTGGGGTGATCCTGCTGGCTGTTGGAGTCTGGGGCAAACCTACTCTGGGCACCTATATCTCCCTTA 150

L D H W G D P A G C W S L G O T Y S G H L Y L P Y

TTGCCGAGAACTCCACAAATGCTCCCTATGTGCTCATCGGAACCTGGCACCCTATTGTTGTCTTTGGCCTGTTTG 225

C R E L H K C S L C A H R N W H H Y C C L W P V W

GATGCTTTGCTACATGTCGTGGTAGCCCATGGATGCTGAAACTGTATGCCATGTTTCTGTCCCTGGTGTTCCTGG 300

H L C Y M S W . P M D A E T V C H V S V P G V P G

CTGAGCTCGTAGCTGGCATTTCAGGGTTTGTGTTTCGTATGAGATCAAGGACACCTTCCTGAGGACTTACACGG 375

. A R S W H F R V C V S S . D O G H L P E D L H G

ACGCTATGCAGACTTACAATGGCAATGATGAGAGGAGCCGGGCAGTGGACCATGTGCAGCGCAGCCTGAGCTGCT 450

R Y A D L O W O . . E E P G S G P C A A O P E L L

GTGGTGTGCAGAACTACACCAACTGGAGCACCAGCCCTACTTCCTGGAGCATGGC 506

W C A E L H O L E H O P L L P G A W

Figure 44

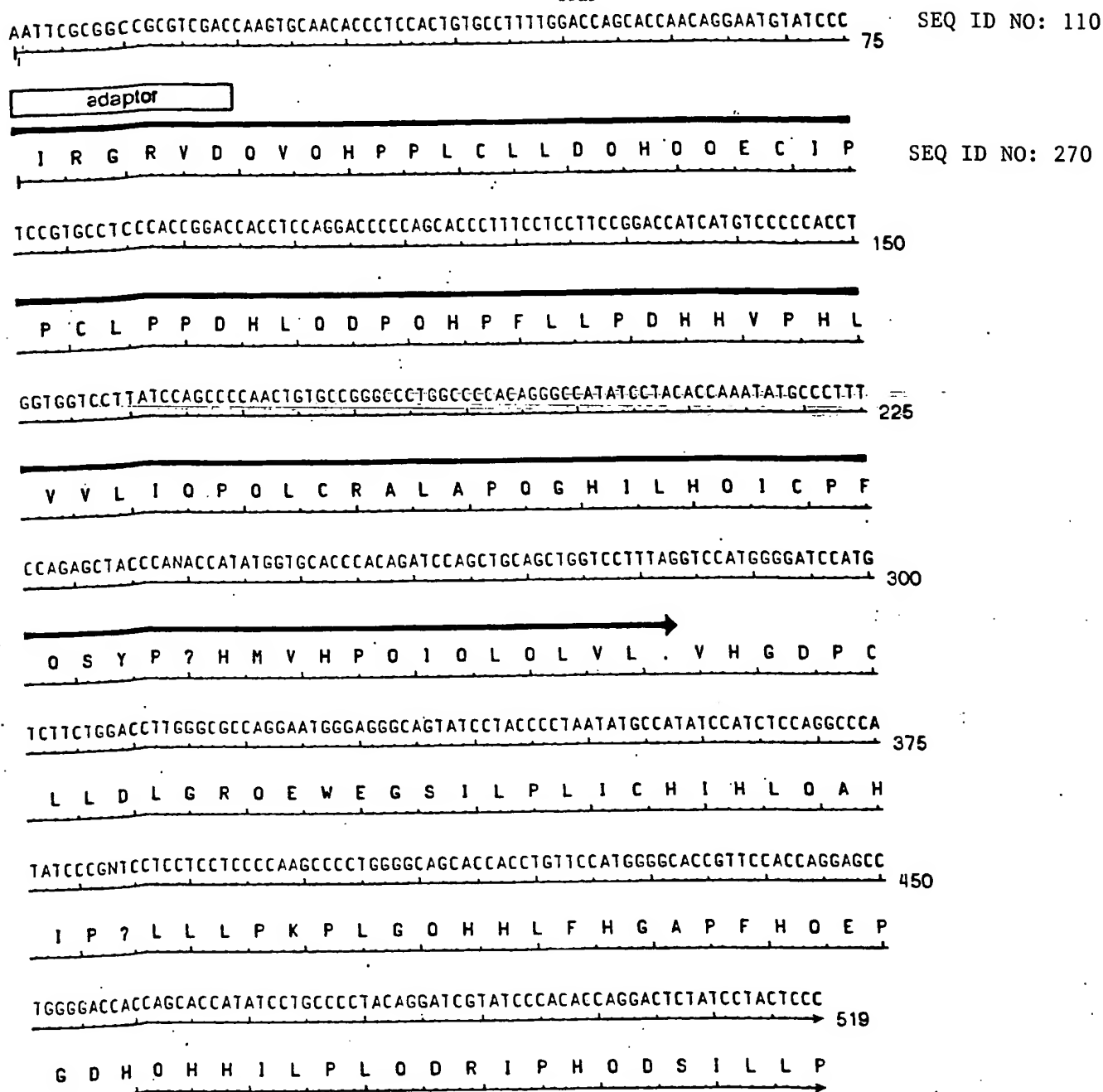


Figure 45

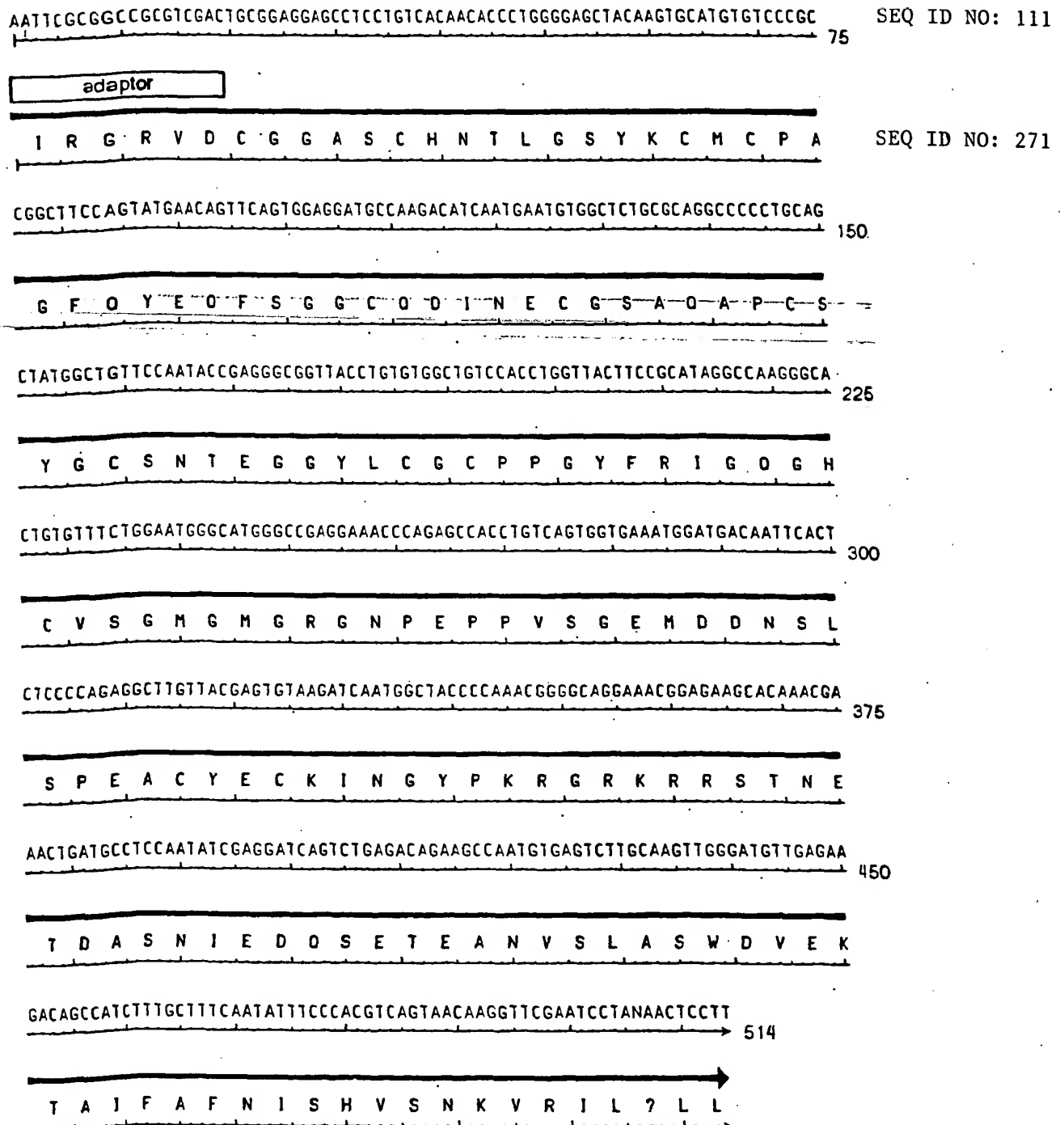


Figure 46

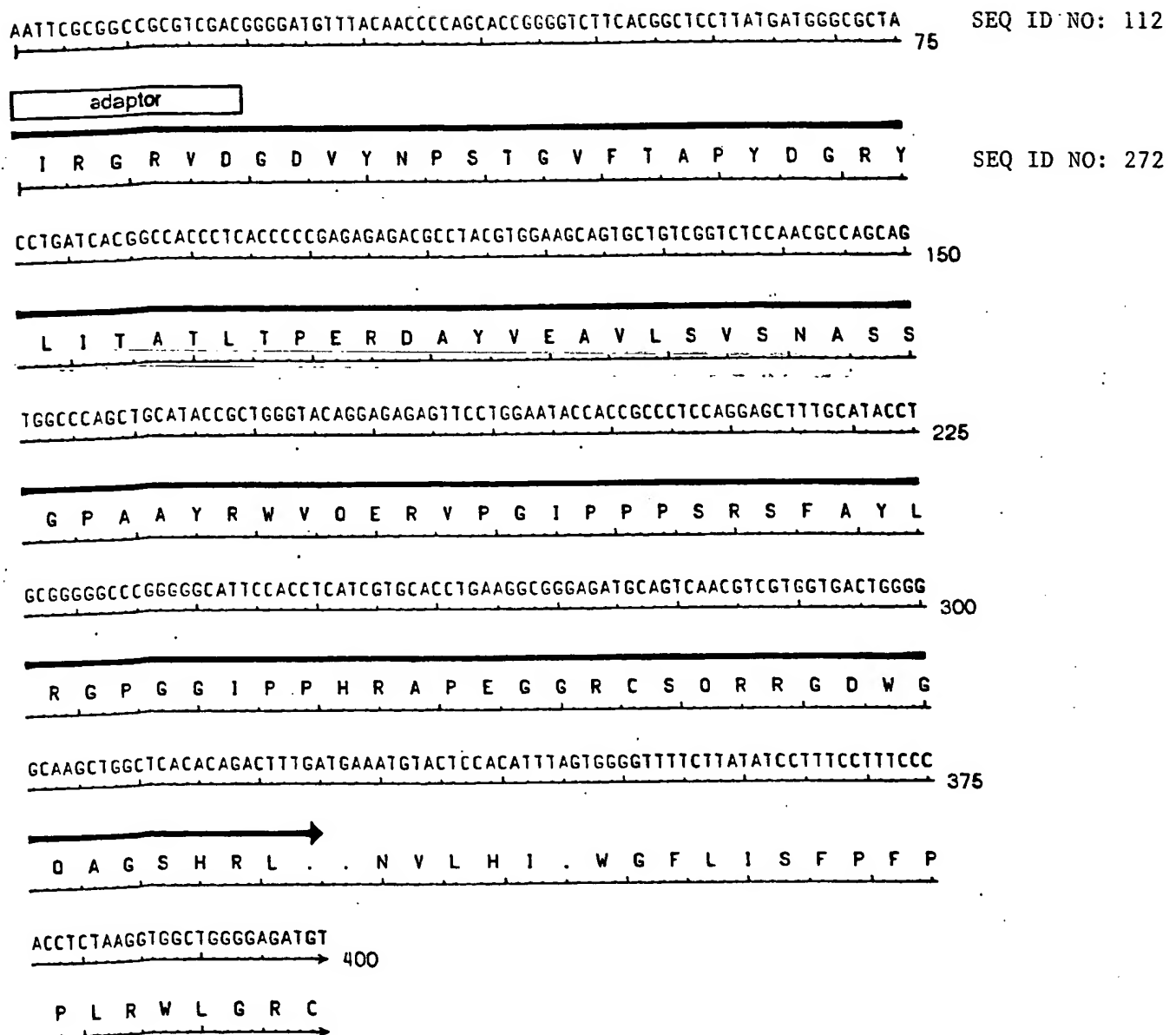
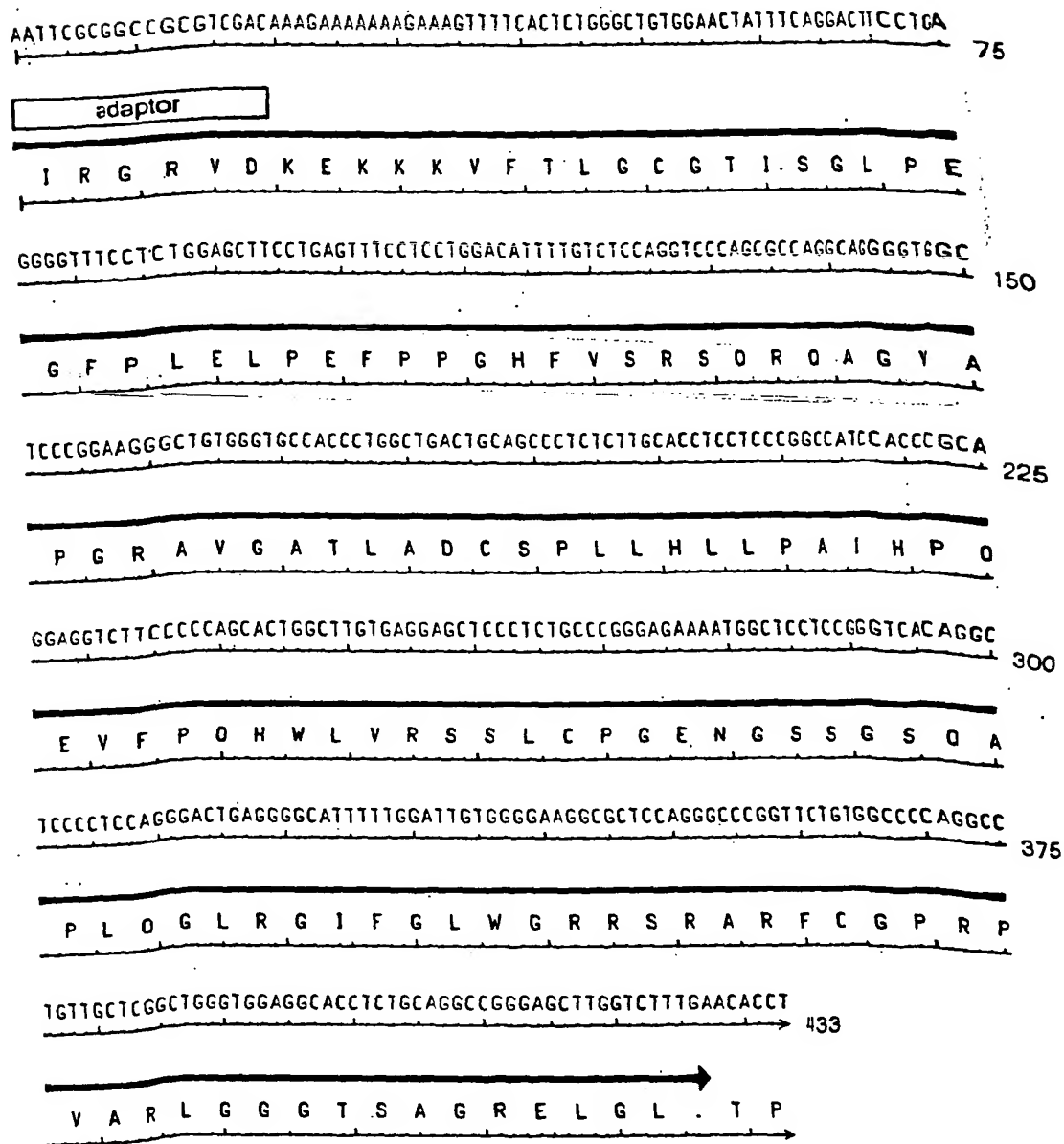


Figure 47



SEQ ID NO: 113

SEQ ID NO: 273

Figure 48

AATTCGCGGCCGCGTCGACGGGGATGTTTACAACCCAGCACCGGGTCTTCACGGCTCCTTATGATGGGCGCTA 75

SEQ ID NO: 114

adaptor

I R G R V D G D V Y N P S T G V F T A P Y D G R Y

SEQ ID NO: 274

CCTGATCACGGCCACCCTCACCCCGAGAGAGACGCTACGTGGAAGCAGTGTCTCGGTCTCCAACGCCAGCAG 150

L I T A T L T P E R D A Y V E A V L S V S N A S S

TGGCCCAGCTGCATACCGCTGGGTACAGGAGAGAGTTCTTGAATACCACCGCCCTCCAGGAGCTTTGCATACCT 225

G P A A Y R W V O E R V P G I P P P S R S F A Y L

CGGGGGGCCCGGGGCATTCCACCTCATCGTGCACCTGAAGGCGGGAGATGCAGTCAACGTCGTGGTGACTGGGG 300

R G P G G I P P H R A P E G G R C S O R R G D W G

GCAAGCTGGCTCACACAGACTTTGATGAAATGTACTCCACATTTAGTGGGGTTTTCTTATATCCTTTCTTTCCC 375

D A G S H R L . . N V L H I . W G F L I S F P F P

ACCTCTAAGGTGGCTGGGGAGATGT 400

P L R W L G R C

Figure 49

AATTGCGGGCCGCGTCGACCGCAACTGTCAAGACATTGATGAGTGTGTGACTGGCATCCACAACCTGCTCCATCAA 75

SEQ ID NO: 115

adaptor

I R G R V D R N C O D I D E C V T G I H N C S I N

SEQ ID NO: 275

CGAGACCTGCTTCAACATCCAGGGCGGCTTCCGCTGCTTGGCCTTCGAGTGCCCTGAGAACTACCGCCGCTCCGC 150

E T C F N I O G G F R C L A F E C P E N Y R R S A

AGCCACGCTCCAGCAGGAGAAGACAGACACGGTCCGCTGCATCAAGTCTGCCGCCCAACGATGTCACATGCGT 225

A T L O O E K T D T V R C I K S C R P N D V T C V

GTTGACCCCGTGACACCATCTCCACACCGTCATCTCGCTGCCTACCTTCCGCGAGTTCACCCGCCCTGAAGA 300

F D P V H T I S H T V I S L P T F R E F T R P E E

GATCATCTTCTCCGGGCCATCACGCCACCGCATCTGCCAGCCAGGCTAACATCATCTTCGACATCACGGAAGG 375

I I F L R A I T P P H P A S O A N I I F D I T E G

GAACCTGCGGGACTCTTTTGACATCATCAAGCGTTACATGGACGGCATGACCGTGGGTGTCGTGCGCCAGGTGCG 450

N L R D S F D I I K R Y M D G M T V G V V R O V R

GCCCATCGTGGGCCCATTTTCATGCCGTCCTGAAGCTGGAGATGAACTATGTGGTCG 506

P I V G P F H A V L K L E M N Y V V

Figure 50

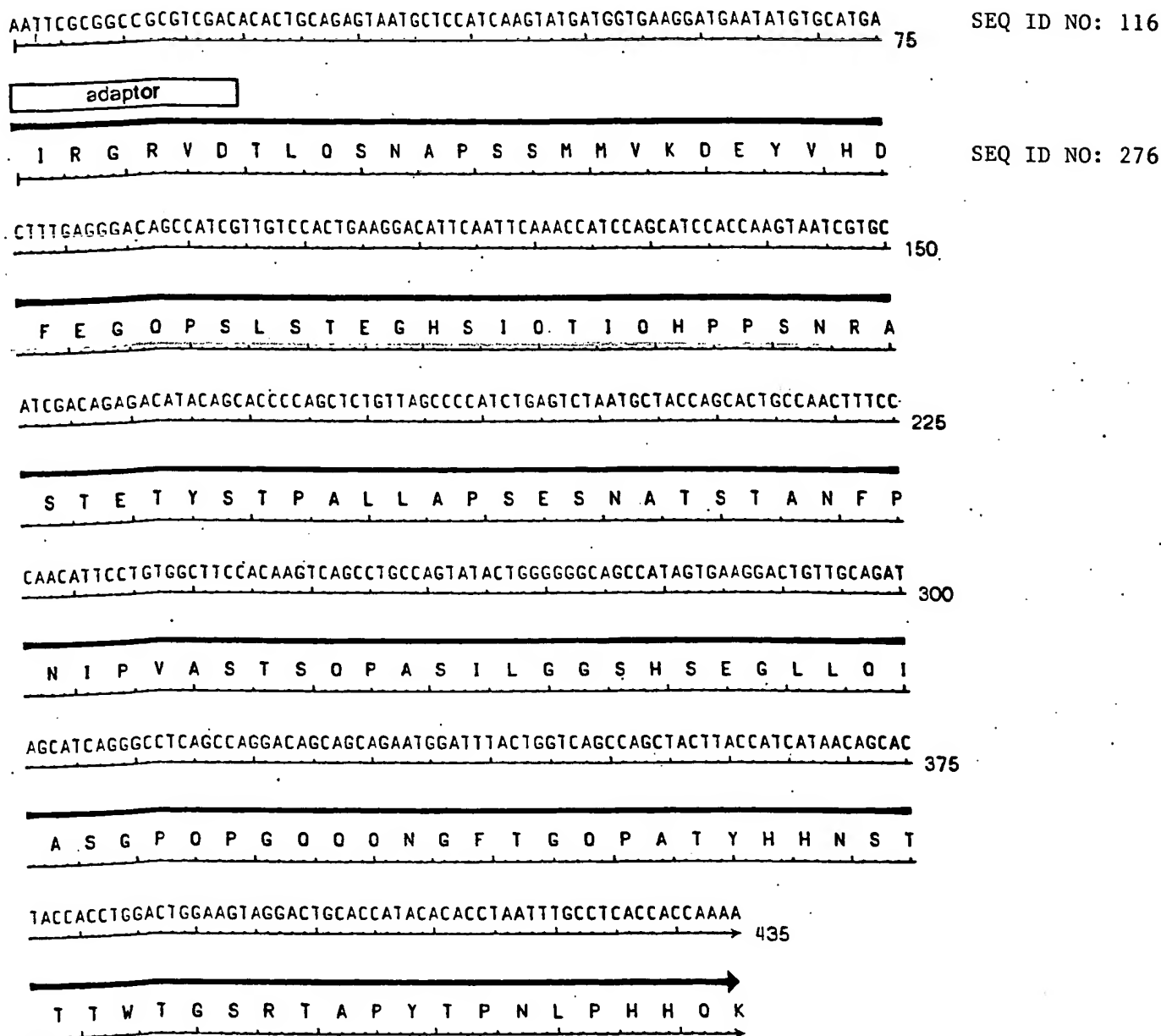
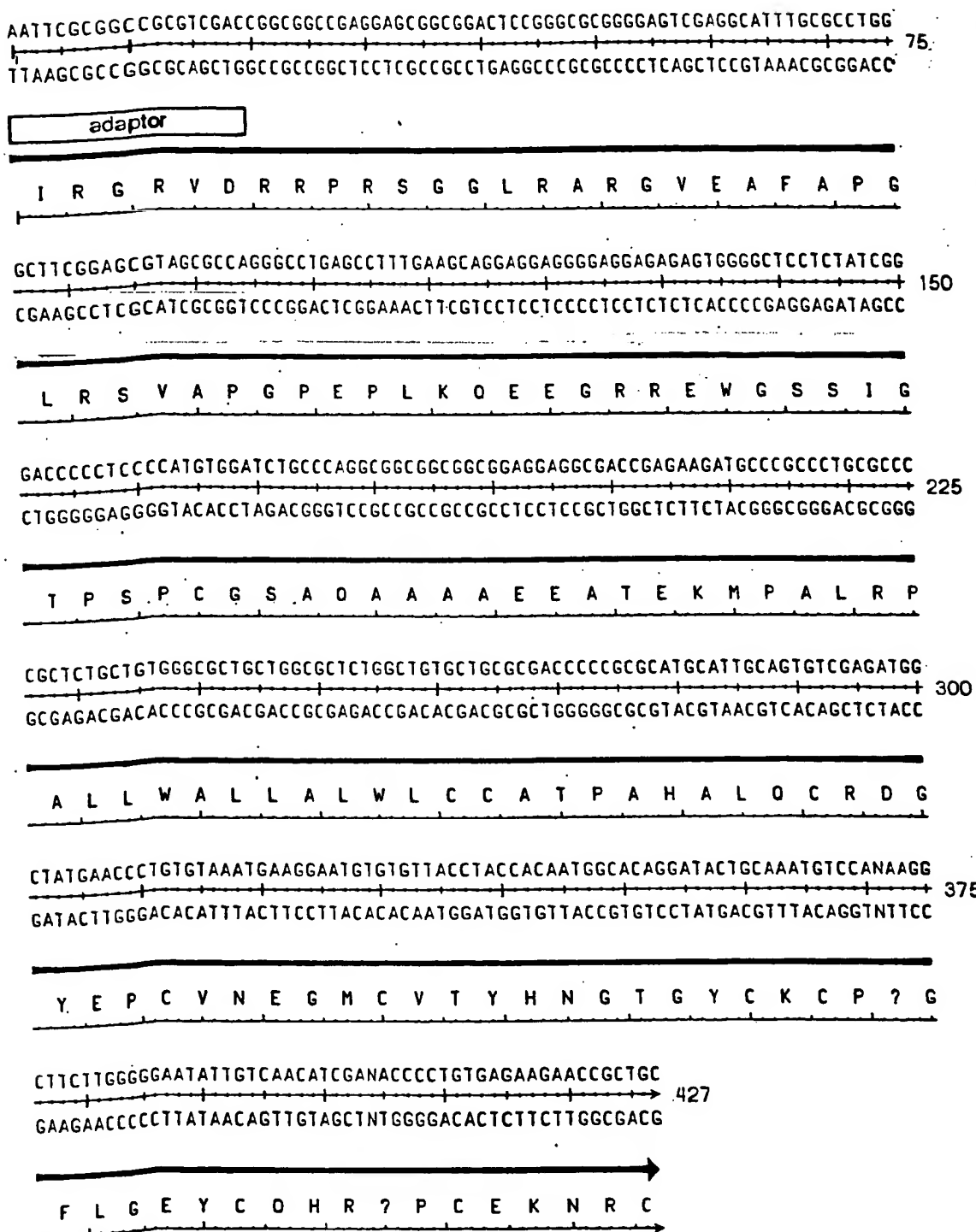


Figure 51



SEQ ID NO: 117

SEQ ID NO: 118

SEQ ID NO: 277

Figure 52

Pathogenicity Island Probe # (bp #)	<i>P. aeruginosa</i> strains that gave Positive signal with the specific probe used	<i>P. aeruginosa</i> strains that gave Negative signal with the specific probe used
3 (25562-26456)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
4 (61181-63607)	PA14, CF2, CF6, CF26, CF29	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF30, CF32
5 (74933-76117)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
6 (84922-86622)	PAO1, PA14, PAO37, CF2, CF, CF26,	PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
7 (103070-104556)	PA14, PAO37, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32
8 (104799-105545)	PA14, CF2, CF6, CF26	PAO1, PAK, CF1, CF3, CF4, CF5, CF27, CF28, CF29, CF30, CF32

These experiments indicate that at least part of the big island region contained in each probe is present in the *P. aeruginosa* strains that gave positive hybridization signal.

Figure 53

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 101 of 118 Customer No.: 21559

ORF7 Protein: SEQ ID NO: 278

MINSHLLYRLSYRGTSFFQPWTLPVLLDSRLRGAPFYGCARACQPSDPKSFSSFSTSDKTALPLHAAALSRLPDAHEKAP
PKRGFPCPPPKRS GEDDLVAFHLRRDTGTRREFAGQDQLRQRVLDPALDGPLQRACAI DRVEADGNQLVQRLLAQFQAQL
ALGQALAQATELDLGDAGDLLASQRLEHHHFVDPVDEFRTFVRI DRVHHCGLRLAVAGQLDLRRTFVGGHHHHGVAEV
HRTPVTVGQASVLEHLEENVEYIRMGLLHLVQQHHRVGLAADRLGQVAAFLEADVARRRADQAGHRVFLHELGHYIPHOR
LLGIEEELGQRLAQGLAHPGRAEEEEERARFVRI GEAGARTAHGVGHGDYRLVLADHSPMQLLLHAQQLLALALEHLRH
RDTGPLGNHFGDFLVGHLVAQQVLVGLAVLVDHLQAAFQVRDGLVLDARHALEVALAPRRLHLLGLLDLLDLRRALHL
GLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFLQLRLALDLQDQATVETIQFLRLGVDLHADAAGGLVDQVDGLV
RQLPIGDVAVRQLGRGDDRAVGDAHPVVHFIAFLEATEDGDGVFLARFVHQLLEAALQRGILLDVLAILEGSSTDAVQ
LAARQSRLEHVAGVHGTFRLAGADHGVQFVDEQDDPAFLLAQFVEDRLQAFLELAELGTGDQRPVHQGQALVLEAVRH
FAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLGDPADLVATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVHR
FAATQVGHGILQRLARHALAEQQLAEPGVLVHRGQQYQLAGDELVALLGQAVSLVEQACEILGQVHVAGRALDLRQVE
FFVEAAQGGDIEADLHQQLDRTALLLEQGGKQVHRLDGRMVMANGQGLGVGERQLQLAGQTVYSHGSSFLL.

Figure 54

Figure 55

clpB protein SEQ ID NO: 279

MRIDRLTSKLQLALSDAQSLAVGHDHPAIEPVHLLSALLEQQGGSIKPLLMQVGFDAALRSGLNKELDALPKIQSPTGD
VNLSQDLARLLNQADRLAQQKGDQFISSSELVLLAAMDENTRLGKLLLGQGVSRKALENAVANLRGGEAVNDPNVEESRQA
LDKYTVDMTKRAEEGKLDPVIGRDDEIRRTIQVLQRRTKNNPVLI GEPGVGKTAIVEGLAQRIINGEVPDGLKDKRLLAL
DMGALIAGAKFRGEFEERLKAVLNELGKQEGRVILFI DELHTMVGAGKAEGAMDAGNMLKPALARGELHCVGATTLDEYR
QYIEKDAALERRFQKVLVDEPSEEDTIAILRGLKERYEVHGV SITDGAI IAAKLSHRYITDRQLPDKAIDLIDEAASR
IRMEIDSKPEELDRDLDRRLIQLKIEREALKKEDDEATRKRRLAKLEEDIVKLEREYADLEEIWKSEKAEVQGSQAQIQKIE
QAKQEMEAARRKGDLESMARIQYQTI PDLERSLQMVDQHGKTENQLLRNKVTDEEIAEVVSKWTGIPVSKMLEGEREKLL
RMEQELHRRVIGQDEAVVAVSNVRRSRAGLADPNRPSGSFLFLGPTGVGKTELCKALAEFLFDTEEALVRIDMSEFMK
HSVARLIGAPPGYVGFEEGGYL TEAIRRKPYSVVLLDEVEKAHPDVFNILLQVLEDGRLTDSHGRTVDFRNTVVVMTSNL
GSAQIQELAGDREAQRAAVMDAVNAHFRPEFINRIDEVVVFEPLAREQIAGIAEIQLGRLRKRLAERELSLELSQEALDK
LIAVGFDVPYGARPLKRAIQRWIENPLAQLILAGKFAPGASISAKVEGDEIVFA.

Figure 56

clpB DNA SEQ ID NO: 120

ATGCGAATAGACCGTTTGACCAGCAAGCTGCAACTGGCGCTCTCCGACGCCAGTCCCTGGCCGTTGGCCATGACCATCC
GGCCATCGAGCCGGTGCACCTGCTTTCCGCCCTGCTCGAGCAGCAAGGCGGTTCGATCAAGCCCCTGCTGATGCAGGTTCG
GCTTCGATATCGCCGCCCTGCGCAGCGGCCCTCAACAAAGAACTCGACGCGCTGCCGAAGATCCAGAGCCCCGACGGCGAC
GTGAACCTGTCCAGGATCTCGCAGCCTGCTCAACCAGGCTGACCGCTGGCCAGCAGAAGGGCGACCAAGTTCATCTC
CAGCGAGCTGGTATTGCTGGCCGCGATGGACGAGAACACCAGGCTCGGCAAGCTGCTGCTCGGCCAGGGCGTGTGCGCGA
AGGGCGTGGAGAATGCCGTGGCCAACCTGCGTGGCGGCGAAGCGGTGAACGACCCGAACGTCGAGGAGTCGCGCCAGGGC
CTGGACAAGTACACCGTCGACATGACCAAGCGCGCCGAGGAAGGCAAGCTCGACCCGGTGATCGGTGCGCAGCAGAGAT
CCGCCGGACCATCCAGGTCTGCGAGCGCGGACCAAGAACAACCCGGTGCTGATCGGCGAACCCGGCGTGGCAAGACCG
CCATCGTCGAGGGCCTGGCCAGCGCATCATCAACGGCGAAGTGCCGGACGGCCTCAAGGACAAGCGCCTGCTGGCCCTG
GACATGGGGCGGTGATCGCCGGTGCCAAGTTCGCGGGCGAGTTCGAGGAACGCCTGAAGGCGGTCTCAACGAAGTGGG
CAAGCAGGAAGGGCGGGTTCATCCTGTTTCATCGACGAAGTGCACACCATGGTTCGGCGCCGGCAAGGCGGAAGGTGCCATGG
ACGCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTGCACTGCGTCGGTGCTACTACCTCGACGAGTATCGC
CAGTACATCGAGAAGGATGCCGCGCTGGAG
CGCCGCTTCCAGAAGGTGCTGGTGACGAACCGAGCGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTA
TGAAGTGCACCACGGGGTGAGCATCACCGACGGCGCGATCATCGCCGCGGCCAAGCTGTCGCACCGCTACATCACCAGT
GGCAACTGCCGGACAAGGCCATCGACCTGATCGACGAGGCGCGCAGCCGATCCGCATGGAGATCGACTCCAAGCCGGAG
GAAGTGGATCGTCTCGACCGTTCGCTGATCCAGCTGAAGATCGAGCGCGAGGCGCTGAAGAAGGAAGACGACGAAGCCAC
CAGGAAGCGCCTGGCCAAGCTGGAGGAGGATATCGTCAAGCTCGAGCGCGAATACGCCGACCTCGAGGAGATCTGGAAGT
CCGAGAAGGCGGAGGTGCAGGGCTCGGCGCAGATCCAGCAGAAGATCGAGCAGGCCAAGCAGGAGATGGAGGCGGCGCGG
CGCAAGGGCGACCTCGAGAGCATGGCGCGCATCCAGTACCAGACCATCCCGGACCTGGAACGCAGCCTGCAGATGGTTCGA
CCAGCACGGCAAGACCGAGAACCAGTTGCTGCGCAACAAGGTGACCGACGAGGAAATCGCCGAAGTGGTTTCCAAGTGGA
CCGGTATCCCGGTGTGCAAGATGCTCGAGGGCGAGCGCGAGAAGCTGCTGCGCATGGAGCAGGAGCTGCATCGGCGAGTG
ATCGGCCAGGACGAGGCGGTAGTCGCCGTGTCCAACGCCGTGCGCCGTTCGCGCGCCGGCCTCGCCGATCCGAACCGGCC
GAGCGGCTCGTTTCTTCTTCTCGGCCCGACCGGGGTGGGCAAGACCGAGTTGTGCAAGGCGCTGGCCGAGTTCCTCTTCG
ATACCGAGGAGGCGCTGGTGCGGATAGATATGTCCGAGTTCATGGAGAACAACACTCGGTGGCCCGCCTGATCGGCGCGCCT
CCGGGCTACGTCCGGCTTCGAGGAAGGCGGC
TACCTGACCGAGGCGATCCGCCGCAAGCCCTACTCGGTGGTGCTGCTGGACGAGGTGGAGAAGGCCCATCCGGATGTATT
CAACATTCTCCTCCAGGTGCTCGAGGACGGACCGCTGACCGACAGTCACGGGCGTACGGTGGACTTCCGCAACACCGTGG
TGGTGATGACCTCCAACCTCGGTTCCGGCGAGATCCAGGAGCTGGCCGGCGACCGCGAGGCGCAACGTGCCGAGTGATG
GACGCGGTCAATGCGCACTTCCGTCCGGAATTTCATCAACCGGATCGACGAAGTGGTGGTGTTCGAGCCGCTGGCTCGCGA
GCAGATCGCCGGCATCGCCGAGATCCAGCTCGGTGCGCTGCGCAAGCGCCTGGCCGAGCGCGAGCTGAGCCTGGAAGTGA
GCCAGGAGGCGCTGGACAAGCTGATTGCCGTGCGCTTCGACCCGGTCTATGGCGCACGCCCGCTGAAGCGGGCCATCCAG
CGCTGGATCGAGAACCGCTGGCGCAACTGATCCTGGCCGGCAATTTCGCGCCGGGTGCCAGTATCTCGGCGAAGGTGGA
AGGGACGAGATCGTCTTCGCTGA

Figure 57

ORF ID	Strand	Left end	Right end	ORF length (aa)	G+C content (%)	Location prediction	Gene name	Gene function	Protein with the highest identity (Gene Name / Species Strain)	E-value (% identity)	GenBank accession no.	Proteins with lesser identity Cut off 30%
RS01	-	801	376	pseudogene	51			Hypothetical protein	RNA-lyase / <i>P. aeruginosa</i> PAO1	8E-37 (86)	AF004531	
RS01	-	959	1,280	pseudogene	57			Hypothetical protein	PA0977 / <i>P. aeruginosa</i> PAO1		AAG04366	
RS02	-	1,247	2,523	1,276	60	cytoplasm	leuG	Integrase	XAG / <i>P. aeruginosa</i> SGLTM	0.00 (94)	AA602084	ISTK4666
RS03	-	2,524	2,901	125	56	cytoplasm		Hypothetical protein	XFI753 / <i>X. fastidiosus</i> 9a5c	3E-34 (49)	AAF84562	XAC2196, XCC3121, STY4665
RS04		2,988	3,850	862	62	cytoplasm		Putative transposase	PA0968 / IS222 / <i>P. aeruginosa</i> PAO1	1E-154 (96)	AA604367	PAI938, RSC2315
RS05	-	3,830	4,138	103	60	cytoplasm		Putative transposase	PA0979 / IS222 / <i>P. aeruginosa</i> PAO1	6E-50 (99)	AA604368	PAI937, RSC2314
RS06		4,190	4,654	464	45	inner or outer membrane		Hypothetical protein	PA0980 / <i>P. aeruginosa</i> PAO1	4E-50 (97)	AA604369	
RS07	+	4,691	5,314	207	46	outer membrane and periplasm		Hypothetical protein	PA0981 / <i>P. aeruginosa</i> PAO1	E-114 (99)	AA604370	
RS08		5,420	5,763	343	50	outer membrane and periplasm		Hypothetical protein	PA0673 / <i>P. aeruginosa</i> PAO1	2E-05 (97)	AA604067	
RS09	+	5,849	6,052	67	57	cytoplasm			No significant similarity			
RS10		6,108	6,326	217	58	cytoplasm			No significant similarity			
RS11	+	6,540	6,989	149	53	inner membrane		Acetyltransferase	PP0651 / <i>P. putida</i> KT2440	2E-75 (95)	AA606276	
RS12		7,418	8,218	800	59	outer membrane and periplasm		Transposase	PP3964 / ISBpu / <i>A.ORE-1</i> / <i>P. putida</i> KT2440	0.00 (98)	AA609558	
RS13	+	8,253	8,822	189	61	cytoplasm		Putative transposase	PA0987 / <i>P. aeruginosa</i> PAO1	4E-88 (85)	AA604376	XAC2424
RS14		8,959	11,022	2,063	59	cytoplasm	ecoliDep2	Gyrolxin (type III secretion system effector)	EX01 / <i>P. aeruginosa</i> PA103	0.00 (100)	AA6038269	
RS15	+	11,019	11,432	137	56	inner membrane	spcU	ExoU chaperone	SpcU / <i>P. aeruginosa</i> PA103	4E-62 (94)	AA616024	

FIGURE 58

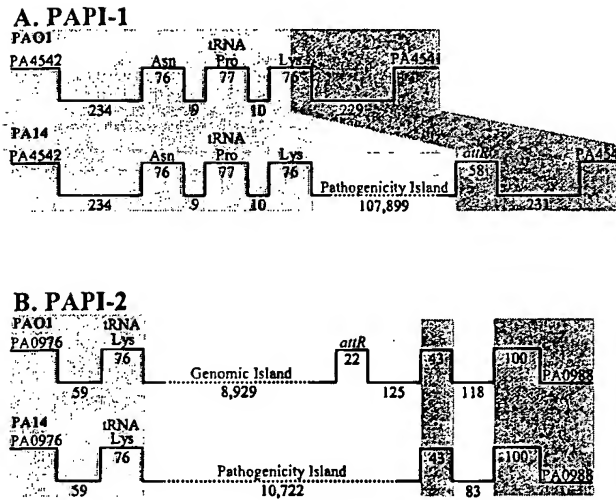


Figure 59

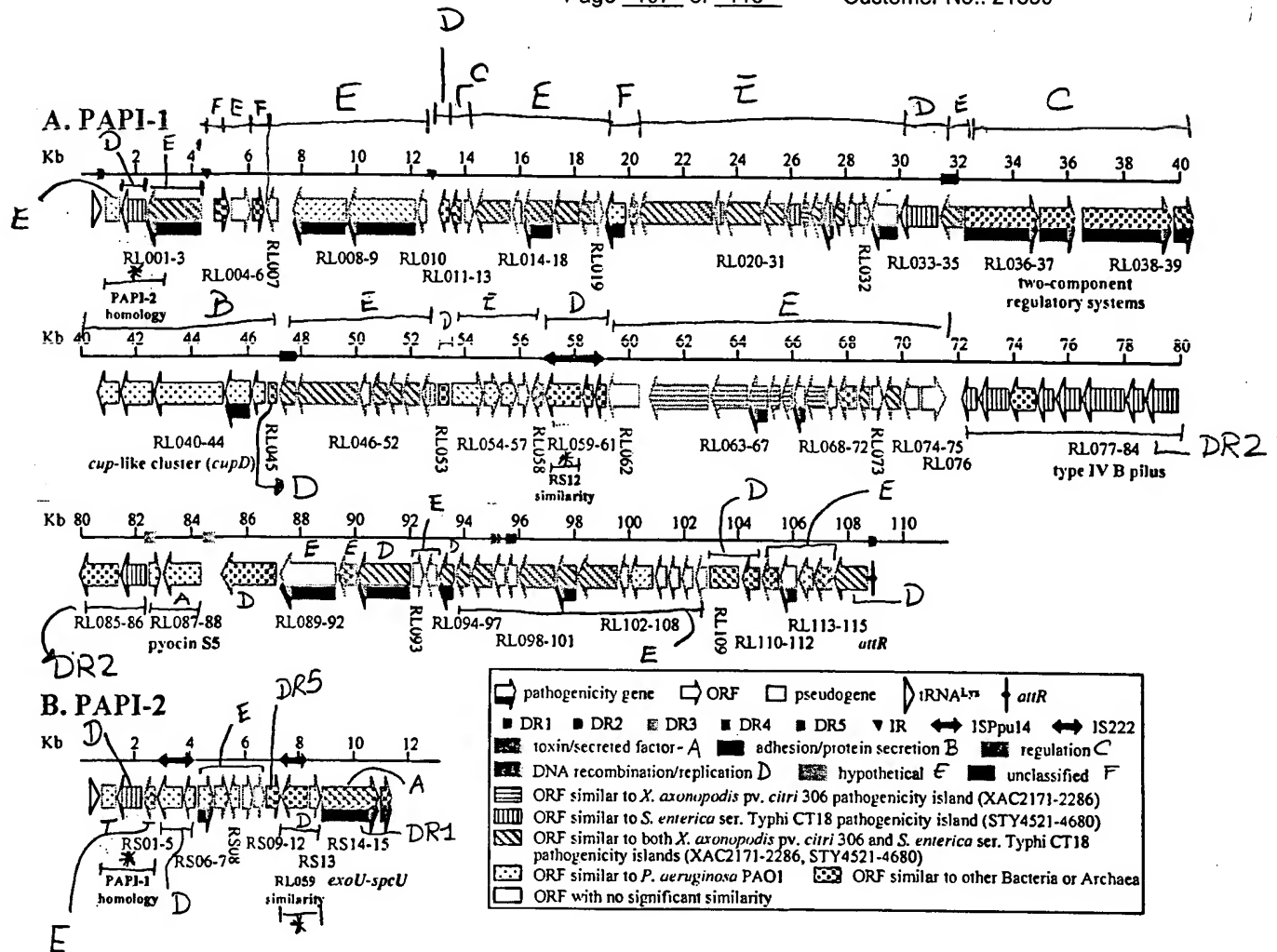
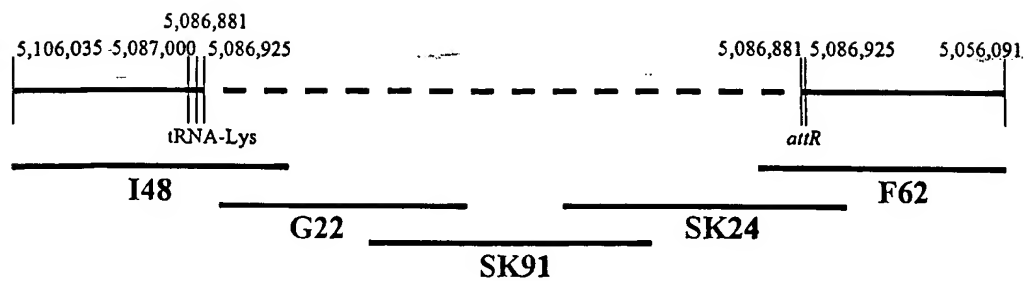


Figure 60

kb	vDR1			vDR2			vDR2			DR3			vDR3			DR1		
	0	10	20	30	40	50	60	70	80	90	100	0	10	20	30	40	50	60
PA14	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
CF2	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
CF6	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
PA037	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
CF26	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
CF29	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
PAK	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++
PAO1	++	+	+	+	+	+	+	+	+	+	++	++	+	+	+	+	+	++

Figure 61

A. PAPI-1



B. PAPI-2

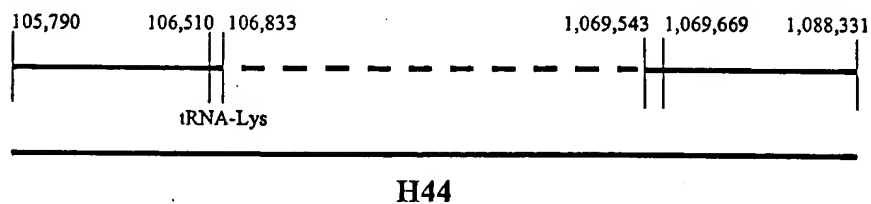


Figure 62

FIGURE 63

Strain name*	% Mouse mortality†	Growth in <i>Arabidopsis</i> leaf‡	Closest published homologue (organism / GenBank accession no.)
PA14	100	4.9×10^6	
RL003 ^b	41	2.3×10^5	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54371
RL008	38	4.1×10^6	<i>M. acetivorans</i> C2A / AAM05538 and <i>P. aeruginosa</i> PAO1 / AAG05323
RL009	31	1.3×10^4	<i>P. aeruginosa</i> PAO1 / AAG05327
RL016	100	2.8×10^4	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54383
RL020	50	3.4×10^5	protein-disulfide isomerase, <i>P. aeruginosa</i> PAO1 / AAG04371
RL022	88	3.3×10^6	<i>P. syringae</i> pv. <i>tomato</i> DC3000 / AAO54394
RL029	38	9.4×10^4	<i>P. aeruginosa</i> C / AAN62148
RL033	25	4.9×10^4	no significant similarity
RL036	44	1.9×10^5	two-component sensor <i>P. aeruginosa</i> PA14 / AAM15532
RL037	43	1.2×10^5	two-component regulator <i>pvrR</i> , <i>P. aeruginosa</i> PA14 AAM15533
RL038	31	4.4×10^4	two-component sensor <i>rscC</i> , <i>S. typhimurium</i> LT2 / AAL21172
RL039	31	2.7×10^5	two-component regulator <i>rscB</i> , <i>E. coli</i> O157:H7 EDL933 / AAG57352
RL043	75	1.7×10^6	probable pili assembly chaperone <i>cupA2</i> , <i>P. aeruginosa</i> PAO1 / AAG05517
RL054	63	NT [§]	<i>P. aeruginosa</i> PAO1 / AAG05610
RL062	78	NT [§]	no significant similarity
RL065	63	4.5×10^5	<i>X. axonopodis</i> pv. <i>citri</i> 306 / AAM37094
RL068	56	2.6×10^5	no significant similarity
RL090	67	2.7×10^4	no significant similarity
RL092	0	1.3×10^5	topoisomerase I TopA, <i>X. fastidiosa</i> 9a5c (plasmid pXF51) / AAF85572
RL095	50	5.3×10^5	single-stranded DNA binding protein Ssb, <i>P. aeruginosa</i> C / AAN62318
RL101	38	1.8×10^6	<i>Pseudomonas</i> sp. B13 / CAD60668
RL112	38	1.6×10^4	no significant similarity
RS06	100	1.8×10^5	<i>P. aeruginosa</i> PAO1 / AAG04369

	Positions	Length (bp)	Number of identical bp	Genes between the repeats
DR1	744-805 108,700-108,762	63	59	PAP1-F (108 Kb)
DR2	31,587-32,248 47,100-47,761	662	654	two component regulatory systems and <i>cup</i> -like cluster (<i>cupD</i>)
DR3	82,574-82,821 85,296-85,540	248	238	pyocin S5 and associated immunity protein
DR4	95,301-95,357 95,358-95,414	57	50	none
DR5	95,767-95,824 95,825-95,881	58	54	none
IR	4,527-4,594 12,825-12,892	68	56	pathogenicity genes and Archaea homologous genes

FIGURE 64

Title: VIRULENCE-ASSOCIATED NUCLEIC ACIDS AND
PROTEINS AND USES THEREOF

Applicants: Laurence Rahme et al.

Filing Date: September 12, 2003 Serial No.: Not Yet Assigned

Page 112 of 118 Customer No.: 21559

IS name	PAPI-1		PAPI-2		Characteristics of IS	
	Position	Length (bp)	Position	Length (bp)	Original length (bp)	IS family
IS _{Ppu14}	56,778-59,119	2,341	7,034-7,999	966	2,383	IS66
IS222	-	-	2,980-4,201	1,222	1,232	IS3

FIGURE 65

FIGURE 66

Function	Prototype name	Type IVB (PAPI-1 in PA14)	Type IVA (PAO1)	Xcp (PAO1)	Hxc (PAO1)	Hpl (PAO1)	Other homologues in PAO1 genome
ATPase	<i>pulE</i>	RL082 (<i>pilO2</i>)	<i>pilB</i> <i>pilT</i> <i>pilU</i>	<i>xcpR</i>	<i>hxcR</i>	<i>hplR</i>	<i>hvbA</i> , <i>hvrA</i>
Peptidase	<i>pulO</i>	RL079 (<i>pilT2</i>)?	<i>pilD</i>	<i>pilD</i>	<i>pilD</i>	<i>pilD?</i>	
Major pilin	<i>pulG</i>	RL080 (<i>pilS2</i>)	<i>pilA</i>	<i>xcpT</i>	<i>hxcT</i>	<i>hplT</i>	
Minor pilin	<i>pulH</i>	RL077 (<i>pilM2</i>)	<i>pilE</i>	<i>xcpU</i>	<i>hxcU</i>	<i>hplU</i>	
	<i>pulI</i>	RL078 (<i>pilV2</i>)	<i>fimU</i>	<i>xcpV</i>	<i>hxcV</i>	<i>hplV</i>	
	<i>pulJ</i>	RL083 (<i>pilP2</i>)	<i>fimT</i>	<i>xcpW</i>	<i>hxcW</i>	<i>hplW</i>	
	<i>pulK</i>	RL086 (<i>pilL2</i>)		<i>xcpX</i>	<i>hxcX</i>	<i>hplX</i>	
Inner membrane protein	<i>pulR</i>	RL081 (<i>pilR2</i>)	<i>pilC</i>	<i>xcpS</i>	<i>hxcS</i>	<i>hplS</i>	<i>xqhA</i>
	<i>pulC</i>	RL084 (<i>pilO2</i>)		<i>xcpP</i>	<i>hxcP</i>		
	<i>pulL</i>			<i>xcpY</i>	<i>hxcY</i>		
	<i>pulM</i>			<i>xcpZ</i>	<i>hxcZ</i>		
Secretin	<i>pulD</i>	RL085 (<i>pilN2</i>)	<i>pilQ</i>	<i>xcpQ</i>	<i>hxcQ</i>		<i>xqhA</i> , <i>xqhB</i> , <i>xqhC</i>

FIGURE 67

ORF 7 (SEQ ID NO: 280)

LEFGSATWTRDPMINSHLLYRLSYRGTSFFQPWTLFVLLDSRLRGAPFYGCARACQPSDPKSFSSPSTSDKTALPLHAAALSRLPDAHEKAPPKR
5 GFPPCPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRLVDPALDGPQLRACAI DRV EADGNQLVQRLLAQFQAQLALGQALAQATELDLGDAGDL
LASQRLHHHFVDPVDEFRTFVRIDRVHHCGLRLAVAGQLDLRRTEVGGHHHGVAEVHRTPTVTGQASVLEHLEENVEYIRMGLLHLVQQHHRV
GLAADRLGQVAAFLEADVARRRADQAGHRVFLHELGHYYPHQRLLGIEELGQRLAQLGLAHPGRAEEEEERAAFPVRI GEAGARTAHGVGHG DYRLV
LADHSPMQLLLHAQQLLALALEHLRHRDTGPLGNHFGDFLVGHVLAQQLVLGLAVLVDHLQAAFPVRDGLVLDARHALEVALAPRRLHLLGLDLL
LDLRRALHLGLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFLLQRLALDLQLDQATVETIQFLRLGVLDLHADAAGGLVDQVDGLVRLPIGDV
10 AVRQLGRGDDRAVGDAHPVVFHIAFLEATEDGDGVFLARFVHQHLEAALQRGILLDLVLAILEVGSSTDAVQLAARQSRLEHVAGVHGTFRLAGADH
GVQFVDEQDDPAFLLAQFVEDRLQAFLELAAELGTGDQRPHVQGGQALVLEAVRHFAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLDPADL
VVATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVRFAATQVGHGILQRLARHALAEQQLAEFGVLVHRGQYQLAGDELVALLGQAVSLVEQA
CEILGQVHVAGRALDLRQRVEFFVEAAQGGDIEADLHQQLDRALTALLEQGGKQVHRLDGRVMANGQGLGVGERQLQLAGQTVYSHGSSFL

ORF7 (SEQ ID NO: 281)

TTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATGATTAAACAGTCATTTGCTCTACCGACTGAGCTATCGCGGAACGCTCTTTCTTCC
AACCTGGAGCGTTCGGTGTGCTGGATTTCGGCTCTCAGAGCGCGCCATTTTACGGATGCGCGCGGCATGTCAACCTCTGATCCAAAAGTTT
TTCTTCTTTTTCACGAGCGACAAAACGGCCCTTCCACTGCATGCGCGAGCGCTCTCGCGCTTACCGGACGCCCATGAAAAAGCCCCCGAAGCGG
20 GGCTTTCCCTGTCCGCCCGAAGAGGTGAGGCGAAGACGATCTCGTCGCTTCCACCTTCGCGGAGATACTGGCACCCGGCGCGAATTTGCCGGCC
AGGATCAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATAGACCGGGTCGAAGCCGACGCAATCAGCT
TGTCAGCGCTCCTGGCTCAGTTCAGGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGCGCATGCCGGCATCTG
CTCGCGAGCCAGCGGCTCGAACACCACTTCTCGATCCGGTGTGATGAATTCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGT
TGCGCTCGCGGTGCGCGGCGAGCTCCTGGATCTGCGCGAACCAGAGGTGAGGTGATCACCACACCGGTGTGCGGAAGTCCACCGTACGCGCGT
25 GACTGTCGGTCAGGCGTCCGCTCCTCGAGCACTGGAGGAGAATGTGAATACATCCGGATGGGCTTCTCCACCTCGTCCAGCAGCACCAACGAGTA
GGGCTTGCGCGGATCGCTCAGGTAGCGGCTTCTCCTCGAGCGGACGTAGCGCGGAGGCGCGCGATCAGGCGGGCCACCGAGTGTCTCTCC
ATGAACCTCGGACATATCTATCCGCACCAAGCGCTCCTCGTATCGAAGAGGAACTCGGCCAGCGCTTGCAAACTCGGTCTTGCCCAACCCCGGTG
GGCCGAGGAAGAGGAACGAGCCGCTCGGCCGTTCCGATCGGCGAGGCGCGCGCAACGGCGCACGGCGTTGGACACGGCGACTACCGCTCGTC
CTGGCCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTG
30 GAAACCACTTCGGCGATTTCTCGTGGTCACTTGTTGCGCAGCACTGGTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCGAGCTGCGTTCCA
GGTCCGGGATGGTCTGGTACTGGATGCGCGCATGCTCTCGAGGTGCGCCCTTGCGCGCGCGCCGCTCCATCTCCTGCTTGGCCTGCTCGATCTTCTG
CTGGATCTGCGCGGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTGCGCGTATTCGCGCTCGAGCTTGACGATATCTCCTCTCC
AGCTTGCGCAGGCGCTTCTGGTGGCTTCGTGCTTCTCTTTCAGCGCTTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCCA
GTTCTCTCCGCTTGAGTCGATCTCCATGCGGATGCGGTGCGCGCTCGTCGATCAGGTGATGGCTTGTCGGCAGTTGCCGATCGGTGATGTA
35 GCGGTGCGACAGCTTGCGCGCGCGATGATCGCGCGTGGTGTGCTACCCCGTGGTGCACTTCATAGCGTTCCTTGAGGCCACGGAGGATGGCG
ATGGTGTCTTCTCGCTCGGTTCTGTCACCAAGCAGCTTCTGGAAGCGGCGCTCCAGCGCGCATCCTTCTCGATGTACTGGCGATACTCGTCGAGGG
TAGTAGCACCGACGAGTCAGCTCGCCGCGCGCAGAGCCGGCTTGAGCATGTTGCCGGCGTCCATGGCACCTTCCGCTTGCCGGCGCGGACCAT
GGTGTGCACTTCGTGATGAACAGGATGACCCGCGCTTCTGCTTGCCAGTTGTTGAGGACCGCTTCAGGCGTTCCTCGAACTCGCCGCGGAAC
TTGGCACCGCGATCAGCGCCCCATGTCCAGGGCCAGCAGGCGCTTGCTCTGAGGCGTCCGCACTTCGCGCTTGATGATGCGCTGGGCCAGGC
40 CCTCGACGATGGCGGTCTTGCCGAGCGCGGTTGCGCGATCAGCACCGGGTGTGTTCTGGTCCGCGCTGCGAGACCTGGATGGTCCGGCGGATCTC
GTCGTGCGACCGATCACCGGTCGAGCTTGCTTCTCGCGCGCTTGCTCATGTGCGAGGTGACTTGTCAGCGCTTGCGCGACTCTCTCGAGC
TTCGGGTGTTTACCGCTTCGCGGCCACGAGGTGGCCACGGCATTCTCAGCGCTTGCGCGACACGCCCTGGCCGAGCAGCTTGCCGAGCC
TGGTGTCTCGTCCATCGCGCCAGCAATACGAGCTCGCTGGAGATGAAGTGGTGCCTTCTGCTGGGCCAGGCGGTGAGCTGGTGGAGCAGGC
TGCGAGATCTGGGACAGGTTACGTGCGCGGTGCGGCTGAGTCTTCGGCAGCGCTGAGTCTTGTGAGGCGCTGCGCAGGCGCGGATA
45 TCGAAGCCGACCTGCATCAGCAGGGGCTTGATCGAACCCTTGTGCTGCGAGCGGCGAAAGCAGGTGCACCGGCTCGATGGCCGATGGTCA
GGCCAACGGCCAGGATGGGCTGCGAGAGCGCAGTTGAGCTTGTGGTCAAACGGTCTATTGATGGTGTGCTCTCTCTCTATAG

FIGURE 68 (SEQ ID NO: 121)

5 ACGTCGGGGGCGCAATTGctACGCCTGcAgAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACCTaGAC
AGGCGGGCTATTCCATATTGACATcACGTCAATGCGGGCCTAATGTTTCGGCCCAGACGGCTGCTAGACAA
GAACCGGCGTAACACCCCTTCCTAGCCTATGCAACTCGCCCCGTAGAAAATGGTGGGTCGTGTAGGATTC
GAACCTACGACCAATTGGTTAAAAGCCAACTGCTCTACCGACTGAGCTAACGACCCAAGTATGAGGTGGT
CGGGGTAGAGAGATTCGAACTCCCGACATCCTGCTCCCAAAGCAGGCGCGCTACCGGACTGCGCTATACC
CCGATTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATG

FIGURE 69

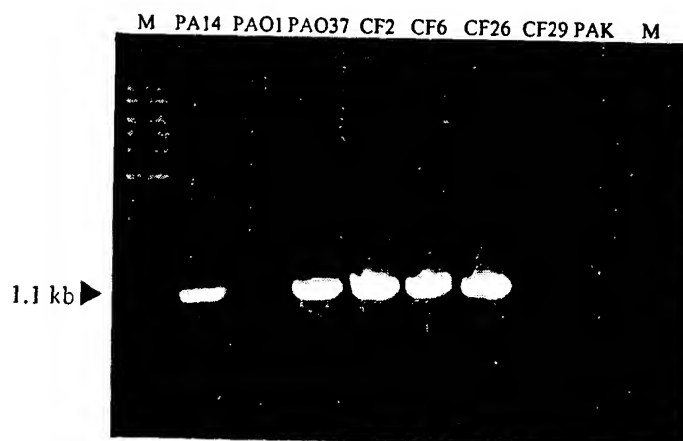


FIGURE 70A

CLUSTAL W (1.82) multiple sequence alignment

```
CF2 (SEQ ID NO: 122) -----GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
PAO37 (SEQ ID NO: 123) -----GATGAAGG-ACCCGAGCGGAACATCCATCTCAC 32
CF6 (SEQ ID NO: 124) -----GATGAAGGACCCGAGCGGAACATCCACCTCAC 33
PA14 (SEQ ID NO: 125) TTTCCGGCACACCCTGGCCACCGACTTGATGAAGGACCCGAGCGGAACATTACCTCAC 60
CF26 (SEQ ID NO: 126) -----GATGAAGG-ACCCGAGCGGAACATCCACCTCAC 32
*****

CF2 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
PAO37 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
CF6 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 93
PA14 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 120
CF26 GAAGTGCCTGCTCAACCACTCGAATATCCAGACCACCATGAGCTACATCGAGGCCGACTA 92
*****

CF2 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
PAO37 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 152
CF6 CGACCACATGCGTGCCGTGCTGCATGCCAGAAGCCTGGCCCAAGGAGCGCTGGAGAACGT 153
PA14 CGATCACATGCGTGCCGTGCTGCATGTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 180
CF26 CGATCACATGCGTGCCGTGCTGCATGTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGT 152
***

CF2 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
PAO37 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
CF6 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 213
PA14 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 240
CF26 CAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCATGCGGGCA 212
*****

CF2 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCGGAGGCCAGGACAGAACCTGCAGAACC 272
PAO37 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCGGAGGCCAGGACAGAACCTGCAGAACC 272
CF6 ACCTCTCGCTCGAATGGGTGAAGTACCGCCGCGGAGGCCAGGACAGAACCTGCAGAACC 273
PA14 ACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAGGCCAGGACAGAGCCTGCAGAACC 300
CF26 ACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAAGCCAGGACAGAGCCTGCAGAACC 272
*****

CF2 AAGGGAGCACATACCAGGGACAGGCATTAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
PAO37 AAGGGAGCACATACCAGGGACAGGCATTAGGGAGGTCCAACCGTGCGGGAAGAAG---C 329
CF6 AAGGGAGCACACACAGGGACAGGCATTAGGGAGGTCCAACCGTGCGGGAAGAAG---C 330
PA14 AAGGGAGCACACGCCAGGGACAGGCATTAGGGAGGTCCAACCGTGCGGGAAGCAGATGC 360
CF26 AAGGGAGCACACACAGGGACAGGCATTAGGGAGGTCCAACCGAGTGGGAAGCAGAGC 332
*****

CF2 GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
PAO37 GCTACCACAGCCACCTGACACCTTCGACCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 389
CF6 GCTACCACAGCCACCTGACACCTTCGATCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 390
PA14 GCTACCACAGCCACCTGACACCTTCGAACCAAAGCGTGCTGTTCACTCTGATGGCTCAAAA 420
CF26 GCTACCACAGCCACCTGACACCTTCGAGCAAAGCGTGCTGTTCACTCTGATGGCTCAACA 392
*****

CF2 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
PAO37 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 449
CF6 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCTGGTGG 450
PA14 CTTATCGAACCGTGCCGCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGATCAGGCG 480
CF26 CTTATCGAACCGTGCCGCTCGGCATCTGCGGCTCCCGCCGCAACCGCGGATCTTGAG 452
*****

CF2 ATGGGGATCTACTGCCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAAGGGCCGGCTA 509
PAO37 ATGGGGATCTACTGCCCCGAAGCAGTCTCGCCTAGCGATACCGATACTGAAGGGCCGGCTA 509
CF6 ATGGGGATCTACCGCCCCGAAGCAGTCTCGCCTAGCGATACCGGTACTGAAGGGCCGGCTA 510
PA14 ATGGGGATCTGCGCCCCGAAGCAATCTCGCCTAGCGATACCGGTACTGA-GGGCCGGCTA 539
CF26 ATGGGGATCTGCGCCCCGAAGCAGCCTCGCCTAGCGATACCGGTACTGAGGGGCCGGCTA 512
*****
```

FIGURE 70B

```
CF2          CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA 569
PAO37        CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAAGAAAAATCTGGAA 569
CF6          CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGTTCGCTAGGCCTGTAGGAAAAATCTGGAA 570
PA14         CCGGACGAAAGGTAGCCGCGCCTCCCAGCAGATCGTTAGGCCTGTAGGAAAAATCTGGAA 599
CF26        CCAGACGAAAGGTAGCCGCGCCTCCCAGCAGATCGCTGGGCCTGTAGGAAAAATCTGGAA 572
             ** ***** **
CF2          TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG 628
PAO37        TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAGTTTCACGG 628
CF6          TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGGGCCC-CGCAATTTCAAGG 629
PA14         TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCAGCGCAATTTCAAGG 659
CF26        TTACCGAGAGCGCCTGGATTCCAGCGCCGGCATGCTGGCAGAGCCC-CGCAATTTTCACGG 631
             ***** **
CF2          CCAAACCGCAGTACCCTCTGTAATCGCTGATTACGTGCGGGGCGCATTGCTACGCCTGC 688
PAO37        CCAAACCGCAGTACCCTCTGTAATCGCTGATTACGTGCGGGGCGCATTGCTACGCCTGC 688
CF6          C-GAAACCGCAGTACCCTCTGTAATCGCTGATTACGTGCGGGGCGCATTGCTACGCCTGC 688
PA14         CCAATACCACAGTACCCTCTGTAATCGCTGATTACGTGCGGGGCGCATTGCTACGCCTGC 719
CF26        C-AAAACCGCAGTACCCTCTGTAATCGCTGATTACGTGCGGGGCGCATTGCTACGCCTGC 690
             * * ** *****
CF2          AGAA-TGGTTTCAGGGCCTTANAACAGAAAAGCCACCTTAAATAGGCGGGCTATT-CC 746
PAO37        AGAAATGGTTTCAGGGCCTTAGAAAACAGAAAAGCCACCTAAATAGGCGGGCTATT-CC 747
CF6          AGAA-TGGTTTCAGAGCCT-GAAAACAGAAAAGNCCACC-TAAATAGGCGGGCTATTTCC 745
PA14         AGAA-TGGTTTCAGGGCCTTAGAAAACAGAAAAGCCACCTAGAAAAGGCGGGCTATT-CC 776
CF26        AGAA-TGGTTTCAGAGCCTTANAACAGAAAAGCCACCTAGATAGGCGGGCTATT-CC 747
             **** *****
CF2          ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCAGACGGCTG--CTGG 798
PAO37        ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCANACGGCTG--CTGG 799
CF6          ATATTGACATCCCG-TCAATGCGGGGCCCTAATGGTTTCGGGCCCANACGGCTTGCTTGG 804
PA14         ATATT-GACATCACG-TCAATGCGGG--CCTAATGTTC--GGCCCAGACGGCTG--CTAG 828
CF26        ATATT-GACATCACGGTCAATGCGGG--GCTAATGTTC-GGGCCCANACGGNTG--CAA 800
             ***** **
```